

Charakterisierung großer SPAST-Deletionen

Dissertation

zur Erlangung des akademischen Grades

doctor medicinae (Dr. med.)

**vorgelegt dem Rat der Medizinischen Fakultät
der Friedrich-Schiller-Universität Jena**

von Claudia Oubrayme

geboren am 21.09.1984 in Nordhausen

Gutachter

- 1. Prof. Dr. Thomas Deufel (Jena)**
- 2. Prof. Dr. Aria Bahnamad (Jena)**
- 3. Prof. Dr. Ludger Schöls (Tübingen)**

Tag der öffentlichen Verteidigung: 07.07.2009

I Abkürzungsverzeichnis

| | |
|-------|---|
| A | Adenin |
| AAA | ATPase assoziiert mit diversen zellulären Aktivitäten |
| Abb | Abbildung |
| ADHSP | autosomal dominante HSP |
| AG | Arbeitsgruppe |
| ATP | Adenosintriphosphat |
| BAC | <i>bacterial artificial chromosome</i> |
| BP | Bruchpunkt |
| BRCA | <i>breast cancer</i> |
| BTK | Bruton Tyrosinkinase |
| C | Cytosin |
| c-DNA | <i>copy-DNA</i> |
| CFTR | <i>cystic fibrosis transmembrane conductance regulator</i> |
| CGH | <i>comparative genomic hybridization</i> |
| CMT1A | Charcot-Marie-Tooth Neuropathie Typ 1A |
| CNA | <i>copy number aberration</i> |
| DMD | Duchenne Muskeldystrophie |
| DNA | Desoxyribonukleinsäure |
| d-NTP | Desoxyribonukleosidtriphosphat |
| ERV | <i>endogenous retrovirus</i> |
| ERV-L | <i>endogenous retrovirus-like</i> |
| f | <i>forward</i> |
| g | Gramm |
| G | Guanin |
| HNPP | <i>hereditary neuropathy with liability to pressure palsies</i> |
| HSP | Hereditäre spastische Paraplegie |
| IDS | Iduronat 2-Sulfatase |
| IGHM | <i>Immunoglobulin Heavy Mu Chain</i> |
| Int | Intron |
| kb | Kilobase |
| LCR | <i>low copy repeat</i> |
| LDL | <i>low density lipoprotein</i> |

| | |
|------|--|
| LINE | <i>long interspersed nuclear element</i> |
| LIS | Lissenzephalie |
| LTR | <i>long terminal repeat</i> |
| MaLR | <i>mammalian LTR retrotransposon</i> |
| MAPH | <i>multiplex amplification and probe hybridization</i> |
| MAR | <i>matrix attachment region</i> |
| Mb | Megabase |
| MEPS | <i>minimal efficient processing segment</i> |
| MER | <i>medium reiterated sequence</i> |
| Mio | Million |
| MIR | <i>mammalian-wide interspersed repeat</i> |
| MLH1 | <i>MutL, E. coli, homolog of 1</i> |
| MLPA | <i>multiplex ligation-dependent probe amplification</i> |
| Mrd | Milliarde |
| MSH2 | <i>MutS, E. coli, homolog of 2</i> |
| NAHR | Nicht allelische homologe Rekombination |
| NF | Neurofibromatose |
| NHEJ | <i>non-homologous end joining</i> (Reparatur von mehrfachen Doppelstrangbrüchen) |
| Nr | Nummer |
| NSD | <i>nuclear receptor binding SET domain protein</i> |
| nt | Nukleotid |
| PAC | <i>P1-derived artificial chromosome</i> |
| PAH | Phenylalaninhydroxylase |
| PCR | Polymerasekettenreaktion |
| PGM | Phosphoglukomutase |
| PHSP | „pure“ (=reine) HSP |
| PKD | <i>polycystic kidney disease</i> |
| PLP | Proteolipid-Protein |
| PMD | <i>Pelizaeus-Merzbacher-Disease</i> |
| PMP | <i>peripheral myelin protein</i> |
| r | reverse |
| SDHC | <i>succinate dehydrogenase 3 integral membrane subunit C</i> |
| sec | Sekunde |

| | |
|------|--|
| SHOX | <i>short stature homeo box</i> |
| SINE | <i>short interspersed nuclear element</i> |
| SMS | Smith-Magenis-Syndrom |
| SNP | <i>single nucleotide polymorphism</i> |
| SPG | <i>spastic paraplegia gene</i> |
| T | Thymin |
| TAP2 | <i>transporter, ATP-binding cassette, major histocompatibility complex 2</i> |
| TSC | <i>tuberous sclerosis</i> |
| UTR | <i>untranslated region</i> |

II Inhaltsverzeichnis

| | |
|--|-----------|
| I Abkürzungsverzeichnis | 3 |
| II Inhaltsverzeichnis | 6 |
| Zusammenfassung | 9 |
| 1. Einleitung | 11 |
| 1.1. Hereditäre spastische Paraplegie | 11 |
| 1.2. SPG4-Mutations-assoziierte hereditäre spastische Paraplegie | 12 |
| 1.3. Große SPAST-Deletionen | 13 |
| 1.4. Verantwortliche Mechanismen für große Deletionen | 15 |
| 1.4.1. Nicht allelische homologe Rekombination | 15 |
| 1.4.2. Reparatur von mehrfachen Doppelstrangbrüchen | 18 |
| 2. Zielstellung | 21 |
| 3. Material und Methoden | 22 |
| 3.1. Geräte und Verbrauchsmaterialien | 22 |
| 3.1.1. Chemikalien | 22 |
| 3.1.2. Geräte | 23 |
| 3.1.3. Verwendete Kits | 24 |
| 3.2. Patienten-DNA | 25 |
| 3.3. Primerdesign und Primer | 26 |
| 3.4. PCR-Screening | 28 |
| 3.5. Charakterisierung der PCR-Produkte | 30 |
| 3.5.1. Agarosegelelektrophorese | 30 |
| 3.5.2. Gelaufreinigung | 31 |
| 3.5.3. Klonierung | 32 |
| 3.5.3.1. Topoisomerasereaktion und Transformation | 32 |
| 3.5.3.2. Kolonie-PCR | 33 |
| 3.5.3.3. Inkubation der Bakterien als Flüssigkultur | 34 |
| 3.5.3.4. Plasmidaufreinigung | 34 |
| 3.5.4. Sequenzierung | 36 |
| 3.6. „Extra-Primer“ zur Validierung | 38 |

| | | |
|-----------|--|-----------|
| 3.7. | In silico-Analyse | 38 |
| 3.7.1. | BLAST | 38 |
| 3.7.2. | Homologie-Analyse | 39 |
| 3.7.3. | Repetitive Sequenzelemente | 39 |
| 3.7.4. | Rekombinations-assoziierte Motive | 40 |
| 3.7.5. | Palindrome | 40 |
| 4. | Ergebnisse | 41 |
| 4.1. | Analyse der bekannten SPAST-Deletionen | 41 |
| 4.1.1. | Deletions-Ausmaß | 41 |
| 4.1.2. | Verteilung der Bruchpunkte auf die Introns | 43 |
| 4.1.3. | Zusammenhang zwischen Anzahl der Bruchpunkte/Intron und Intron-Größe | 44 |
| 4.1.4. | Zusammenhang zwischen dem Auftreten von Bruchpunkten und dem Alu-Gehalt | 45 |
| 4.2. | Klonierung der zehn Fusionssequenzen | 47 |
| 4.2.1. | PCR-Produkte | 47 |
| 4.2.2. | Kolonie-PCR-Produkte | 49 |
| 4.2.3. | Sequenzen | 50 |
| 4.3. | In silico-Analyse der Bruchpunktregionen | 52 |
| 4.3.1. | Homologie an den Fusionspunkten | 52 |
| 4.3.2. | Homologie in den Bruchpunktumgebungen | 52 |
| 4.3.3. | Strukturelle Besonderheiten in den Bruchpunktumgebungen | 53 |
| 4.3.3.1. | Repetitive Sequenzelemente | 53 |
| 4.3.3.2. | Rekombinations-assoziierte Motive | 55 |
| 4.3.3.3. | Palindrome | 56 |
| 4.3.3.4. | AT-Reichtum | 57 |
| 5. | Diskussion | 58 |
| 5.1. | Methode | 58 |
| 5.2. | PCR-Ergebnisse | 61 |
| 5.2.1. | Falsch positive PCR-Ergebnisse | 61 |
| 5.2.2. | Falsch negative PCR-Ergebnisse | 62 |

| | | |
|------------|-----------------------------------|-----------|
| 5.3. | Mechanismus und Struktur | 64 |
| 5.3.1. | Hinweise auf NHEJ | 64 |
| 5.3.2. | Hinweise auf NAHR | 66 |
| 5.3.3. | Dominanz eines Mechanismus | 67 |
| 5.4. | Bruchpunkt-Hotspot | 71 |
| 6. | Schlussfolgerung | 74 |
| III | Literaturverzeichnis | 75 |
| IV | Danksagung | 86 |
| V | Ehrenwörtliche Erklärung | 87 |
| VI | Wissenschaftlicher Beitrag | 88 |

Zusammenfassung

Erbkrankheiten können durch vielerlei Arten von Mutationen verursacht sein. Neben den häufigen Basenaustauschen und Leserahmenverschiebungen gehören dazu auch die so genannten genomischen Umordnungen (*genomic rearrangements*) wie Inversionen, Translokationen, Duplikationen und Deletionen. Besonders Letztere erlangten in den vergangenen Jahren mehr und mehr an Bedeutung. Das liegt nicht an einem gehäuftem Auftreten sondern vielmehr an einer Weiterentwicklung der diagnostischen Technologien.

Die Vorgänge, die in großen genomischen Deletionen resultieren, sind noch wenig erforscht. Ein Hauptgrund dafür dürfte das meist fehlende Wissen um die exakten Ausdehnungen dieser Aberrationen sein: nur ein Bruchteil der bekannten Deletionen ist bis dato auf Sequenzebene charakterisiert. Die wenigen derartigen Studien lassen auf zwei prinzipielle Mechanismen schließen. Dabei handelt es sich um NHEJ (*non-homologous end joining* = Reparatur von mehrfachen Doppelstrangbrüchen) und um NAHR (Nicht allelische homologe Rekombination). NHEJ stellt einen Reparaturmechanismus von benachbarten Doppelstrangbrüchen dar, wobei die Sequenz zwischen den Bruchpunkten verloren geht und damit eine Deletion entsteht. Voraussetzung für NAHR ist eine Zusammenlagerung von nicht allelischen, aber hoch homologen DNA-Abschnitten. Deletionen können resultieren, wenn diese illegitimen Zusammenlagerungen zum Ausschneiden von Sequenzen zwischen den homologen DNA-Abschnitten führen.

Große genomische Deletionen sind seit Kurzem auch für das SPAST-Gen beschrieben. Mutationen in diesem Gen verursachen die hereditäre spastische Paraplegie (HSP) vom Typ 4 (SPG4), eine autosomal dominant vererbte Bewegungsstörung. Klinisches Hauptmerkmal ist eine progrediente Spastik und Schwäche im Bereich der unteren Extremitäten. Pathologisch wird eine Schädigung des Tractus corticospinalis beobachtet.

In der vorliegenden Arbeit wurde zum einen untersucht, ob im SPAST-Gen, wie in vielen anderen Genen, auch eine Bruchpunkt-Häufung nachweisbar ist. Dazu wurden alle 63 bekannten Deletionen hinsichtlich der Verteilung ihrer Bruchpunkte untersucht. Zum anderen wurde versucht, von einer Auswahl zehn intragenischer SPAST-Deletionen die Fusionssequenzen zu ermitteln, um auf den Mechanismus schließen zu können. Die Lösung dieser beiden Aufgaben basierte auf Literaturrecherchen, bioinformatischen Analysen und statistischen Berechnungen, aber auch auf PCR-Reaktionen, Klonierungen und Sequenzierungen.

Die Analyse der Bruchpunkt-Verteilung zeigte, dass in fast allen SPAST-Introns Bruchpunkte zu finden sind. Weiterhin besteht eine starke Korrelation zwischen Bruchpunkt-Anzahl und Intron-Größe, nicht aber zwischen Bruchpunktdichte und Gehalt an repetitiven Alu-Elementen. Die daraus ableitbare weite Verteilung der Bruchpunkte lässt auf das Fehlen eines Hotspots schließen. Von acht der zehn ausgewählten Deletionen gelang die Amplifikation und Sequenzierung der Fusionssequenzen. Zwei davon weisen eine ausgeprägte Homologie im Bereich ihrer Bruchpunktumgebungen auf. Die Bruchpunkte finden sich in homologen Bereichen parallel orientierter Alu-Elemente. Diese Aberrationen können somit durch NAHR erklärt werden. Dagegen sind die anderen sechs Fusionen gekennzeichnet durch eine Mikrohomologie von 3-7nt im Fusionsbereich und eine geringe Übereinstimmung zwischen 5'- und 3'-Bruchpunktumgebung. Dies spricht für NHEJ als verantwortlichen Mechanismus. Allerdings sind auch bei diesen Ereignissen einzelne Bruchpunkte in repetitiven Elementen lokalisiert. Zusätzlich wurden in den Bruchpunktumgebungen aller Fusionen Motive gefunden, die von anderen Studien mit dem Auftreten von Deletionen in Verbindung gebracht wurden (*Immunoglobulin-Heavy-Chain-Class-Switch-Repeats*, AT-reiche Nukleotidabfolgen, Translin-Bindestellen, Palindrome). Diese Sequenzen sind allerdings nicht überrepräsentiert im Vergleich zur restlichen intronischen SPAST-Sequenz.

Das Fehlen eines Hotspots und einer starken Assoziation mit repetitiven Elementen sowie das Überwiegen von NHEJ sind gut mit einer zufälligen Verteilung von SPAST-Deletionen vereinbar. Die Ergebnisse sind insofern von diagnostischer Relevanz, dass sie die Idee einer Fokussierung der Deletionsanalyse auf bestimmte Exons zurückweisen. Die acht ermittelten Fusionssequenzen stellen eine Bereicherung des bislang kleinen Datensatzes zu Deletionen dar. Die hier durchgeführten Analysen zu potentiellen Deletions-fördernden Motiven zeigen, dass eine noch viel umfangreichere Erweiterung für statistisch robuste Aussagen unersetzlich ist. Der Vergleich von SPAST mit anderen Genen ergibt sowohl Unterschiede als auch Gemeinsamkeiten. Die vergleichenden Betrachtungen stellen damit möglicherweise auch eine Basis für eine Klassifizierung von Deletions-anfälligen Genen dar.

1. Einleitung

1.1. Hereditäre spastische Paraplegie

Bei der hereditären spastischen Paraplegie (HSP) handelt es sich um eine Gruppe klinisch und genetisch stark heterogener neurodegenerativer Erkrankungen (Fink 2003, Harding 1983). Hauptsymptome sind eine bilaterale progressive Spastik und Schwäche der unteren Extremitäten, wobei eine Diskrepanz zwischen der Stärke der Tonussteigerung und der Parese typisch ist (McDermott et al. 2000, Reid 1999, Reid 2003). Außer Gangstörung, Pyramidenbahnzeichen, Hyperreflexie im Bereich der unteren und oberen Extremitäten und gestörter Vibrationsempfindung, ist auch eine Störung der Blasenfunktion im Sinne einer Dranginkontinenz häufig (Harding 1981). Der Schweregrad bewegt sich zwischen subklinischer Betroffenheit (10-20%) bis hin zur starken Behinderung mit Rollstuhlpflicht. Allerdings ist der Verlauf durch eine langsame Progredienz gekennzeichnet (Reid 1999, Harding 1981). Der Krankheitsbeginn variiert von früher Kindheit bis zum mittleren Erwachsenenalter, wobei ein früher Beginn mit langsamer Progredienz und eine späte Manifestation mit schnellem Fortschreiten der Erkrankung einhergehen (Harding 1981, Harding 1983, Reid 1999).

Neuropathologisches Korrelat ist eine axonale Degeneration, welche die terminalen Enden des Tractus corticospinalis und des Hinterstranges sowie in geringerem Ausmaß den aufsteigenden Tractus spinocerebellaris betrifft (Reid 1999, Fonknechten et al. 2000, McDermott et al. 2000).

Klinisch wird die reine HSP von der komplizierten Form unterschieden. Zusätzliche Symptome der komplizierten HSP können cerebelläre Ataxie, retinale Degeneration, kognitive Beeinträchtigung, Demenz oder Epilepsie sein (Reid 1999, Harding 1983).

Die reine HSP ist die größte der beiden Subgruppen mit einer Prävalenz von bis zu 9,6/100000, abhängig von der betrachteten Population (Polo et al. 1991). Sowohl die reine als auch die komplizierte hereditäre spastische Paraplegie können autosomal dominant, autosomal rezessiv oder durch X-chromosomalen Erbgang weitergegeben werden (Stevanin et al. 2007, Hanein et al. 2008, Keppen et al. 1987). Dabei stellt die autosomal dominante HSP (ADHSP) den größten Anteil (70-80%) der reinen hereditären spastischen Paraplegie dar (Reid 1999, Harding 1983). Insgesamt wurden für alle drei Vererbungsmuster 37 assoziierte Loci identifiziert (Hanein et al. 2007). Bekannt sind derzeit 16 Gene (Fink 2003, Mannan et al. 2006, Valdmann et al. 2007, Stevanin et al. 2007, Tsaousidou 2008).

Das Hauptkrankheits-Gen ist SPAST (SPG4-Lokus), dessen Mutationen für über 50% der ADHSP und für ca. 40% aller Fälle von hereditärer spastischer Paraplegie verantwortlich sind (Yabe et al. 2002, Fonknechten et al. 2000, Sauter et al. 2002).

1.2. SPG4-Mutations-assoziierte hereditäre spastische Paraplegie

SPAST liegt auf dem Chromosom 2p21-2p22, besteht aus 17 Exons, ist ca. 90kb lang und kodiert für Spastin, ein Protein der AAA-Protein-Familie (*ATPases associated with diverse cellular activities*) (Hazan et al. 1999, McDermott et al. 2000). Spastin ist mit dem Mikrotubuli-Zytoskelett assoziiert und bindet die Mikrotubuli am N-terminalen Ende. Die Interaktion ist transient und ATP-abhängig. Somit scheint Spastin involviert zu sein in die Mikrotubuli-Dynamik (Errico et al. 2002).

Klinisch manifestieren sich SPAST-Mutationen fast ausschließlich mit einem reinen, meist milden Phänotyp. Dies bedeutet eine spastische Paraparese, welche assoziiert ist mit einer Hyperreflexie, verminderter Vibrationsempfindung an den unteren Extremitäten, Fußdeformitäten und Urin-Inkontinenz (Lindsey et al. 2000, McDermott et al. 2006). Das Alter bei Krankheitsbeginn ist stark variabel und reicht von früher Kindheit bis in die achte Dekade, wobei das mittlere Alter zu Beginn der Erkrankung in verschiedenen Veröffentlichungen zwischen zwölf und 43 Jahren schwankt (Depienne et al. 2007, McDermott et al. 2000). So liegt bei Mc Monagle et al. 2000 der Gipfel in der vierten Dekade und ist auch in anderen Veröffentlichungen ähnlich (Iwanaga et al. 2005, Yabe et al. 2002, Yip et al. 2003). Insgesamt ist der Krankheitsbeginn bei SPAST-Mutationen jedoch später im Vergleich zu anderen HSP-Formen, die auf Veränderungen im SPG3, SPG5 oder SPG7 zurückzuführen sind (McMonagle et al. 2000, Fonknechten et al. 2000). Außerdem wird eine altersabhängige und unvollständige Penetranz berichtet (McMonagle et al. 2000, Fonknechten et al. 2000). Es besteht eine starke klinische Variabilität hinsichtlich des Alters bei Krankheitsbeginn, Schweregrad der Symptomatik und Progression, selbst innerhalb der Familien (Starling et al. 2002). Es wurden sogar Tendenzen berichtet, dass Kinder von ebenfalls betroffenen Eltern früher und stärker mit Symptomen auffällig wurden als ihre Eltern. Daher wurde zunächst eine CAG-Repeat-Expansion als Begründung dafür vermutet (Nielsen et al. 1997). Diese konnte aber bei McMonagle et al. 2000 durch Gen-Klonierung ausgeschlossen werden.

Bis heute wurden über 150 verschiedene Mutationen gefunden (Depienne et al. 2006, Patrono et al. 2005, Fonknechten et al. 2000, Erichsen et al. 2007, Svenson et al. 2001 und weitere kleine Studien), welche die Hypothese, die CAG-Repeat-Expansion sei eine häufige Ursache für Antizipation, widerlegen (McMonagle et al. 2000). Dabei handelt es sich hauptsächlich um Missense-, Nonsense- und Splicestellen-Punktmutationen bzw. Leserahmen-Mutationen wie kleine Deletionen und Insertionen (McMonagle et al. 2000, Svenson et al. 2001, Yabe et al. 2002, Meijer et al. 2002, Burger et al. 2000). Diese betreffen fast alle Exons (Meijer et al. 2002, Yabe et al. 2002, Svenson et al. 2001, Burger et al. 2000, Lindsey et al. 2000).

1.3. Große SPAST-Deletionen

Bei einem beträchtlichen Anteil der Patienten, bei denen eine SPG4-Assoziation gesichert war, wurden keine kleinen Mutationen gefunden (Svenson et al. 2001, Starling et al. 2002, Yabe et al. 2002, Lindsey et al. 2000, McMonagle et al. 2000). Daraus wurde geschlussfolgert, dass es sich hierbei entweder um Mutationen im Promotorbereich handelt oder multiexonische Kopiezahlanomalien vorliegen. Beide können mit einem konventionellen PCR-basierten Mutations-Screening nicht entdeckt werden (McMonagle et al. 2000, Svenson et al. 2001, Meijer et al. 2002). Schon Burger et al. gaben 2000 an, dass eine Deletion kompletter Exons möglich sei bei Punktmutations-negativen ADHSP-Patienten. Sauter et al. fanden 2002 eine große SPAST-Deletion (DelEx13_16) bei der Untersuchung von 149 HSP-Familien und Iwanaga et al. identifizierten 2005 mittels spezifischer m-RNA-Expressions-Studie eine 2kb große Deletion im Bereich der 5'UTR und des Intron1 im SPAST-Gen.

Aber erst Beetz et al. konnten 2006 mit einem selbst entwickelten MLPA-Assay eine beträchtliche Anzahl partieller SPAST-Deletionen identifizieren. So wurden bei 65 Punktmutations-negativen HSP-Patienten, welche eine Assoziation zu SPG4 aufwiesen, zwölf Fälle (18%) mit großen Deletionen gefunden, womit diese Art der Aberration einen hohen Anteil aller SPAST-Mutationen ausmacht (Beetz et al. 2006). Der Publikation von Beetz et al. 2006 folgte die Veröffentlichung von 24 partiellen SPAST-Deletionen bei 121 Punktmutations-negativen HSP-Patienten durch Depienne et al 2007, die mittels desselben MLPA-Kits identifiziert werden konnten. Klinisch stellte sich das Krankheitsbild der HSP bei Depienne et al. 2007, genauso wie bei Beetz et al. 2006, dem der Punktmutations-positiven Patienten ähnlich dar. Nach sieben weiteren durch Beetz et al 2007 publizierten großen Deletionen, konnten auch durch Erichsen et al. weitere vier partielle SPAST-Deletionen 2007

identifiziert werden. Es folgten weitere Einsendungen (siehe Tabelle 1.1) an die AG Klinische Chemie (Jena), womit insgesamt 31 solcher Aberrationen hier identifiziert werden konnten. Tabelle 1.1 zeigt alle bisher bekannten partiellen SPAST-Deletionen mit den jeweils deletierten Exon-Kombinationen.

| Laufende Nr. | Deletion | Quelle |
|--------------|----------|-------------------------|
| 1 | Ex1 | Beetz 06 Neurology |
| 2 | Ex1_3 | Beetz 06 Neurology |
| 3 | Ex1_7 | Beetz 06 Neurology |
| 4 | Ex2_5 | Beetz 06 Neurology |
| 5 | Ex5_7 | Beetz 06 Neurology |
| 6 | Ex5_15 | Beetz 06 Neurology |
| 7 | Ex5_15 | Beetz 06 Neurology |
| 8 | Ex6_17 | Beetz 06 Neurology |
| 9 | Ex10 | Beetz 06 Neurology |
| 10 | Ex10_12 | Beetz 06 Neurology |
| 11 | Ex13_16 | Beetz 06 Neurology |
| 12 | Ex16_17 | Beetz 06 Neurology |
| 13 | Ex17 | Beetz 06 Neurology |
| 14 | Ex1 | Depienne 07 J Med Genet |
| 15 | Ex1 | Depienne 07 J Med Genet |
| 16 | Ex1 | Depienne 07 J Med Genet |
| 17 | Ex1_4 | Depienne 07 J Med Genet |
| 18 | Ex1_17 | Depienne 07 J Med Genet |
| 19 | Ex1_17 | Depienne 07 J Med Genet |
| 20 | Ex1_17 | Depienne 07 J Med Genet |
| 21 | Ex1_17 | Depienne 07 J Med Genet |
| 22 | Ex4_17 | Depienne 07 J Med Genet |
| 23 | Ex5_6 | Depienne 07 J Med Genet |
| 24 | Ex5_7 | Depienne 07 J Med Genet |
| 25 | Ex6 | Depienne 07 J Med Genet |
| 26 | Ex8_12 | Depienne 07 J Med Genet |
| 27 | Ex8_17 | Depienne 07 J Med Genet |
| 28 | Ex8_17 | Depienne 07 J Med Genet |
| 29 | Ex9 | Depienne 07 J Med Genet |
| 30 | Ex9_12 | Depienne 07 J Med Genet |
| 31 | Ex10_16 | Depienne 07 J Med Genet |
| 32 | Ex13 | Depienne 07 J Med Genet |
| 33 | Ex13 | Depienne 07 J Med Genet |
| 34 | Ex16 | Depienne 07 J Med Genet |
| 35 | Ex16 | Depienne 07 J Med Genet |
| 36 | Ex16_17 | Depienne 07 J Med Genet |
| 37 | Ex17 | Depienne 07 J Med Genet |
| 38 | Ex2_16 | Beetz 07 Hum Mut |
| 39 | Ex4_17 | Beetz 07 Hum Mut |
| 40 | Ex10_12 | Beetz 07 Hum Mut |
| 41 | Ex17 | Beetz 07 Hum Mut |
| 42 | Ex16 | Beetz 07 Hum Mut |
| 43 | Ex8_17 | Beetz 07 Hum Mut |
| 44 | Ex17 | Beetz 07 Hum Mut |

| Laufende Nr. | Deletion | Quelle |
|--------------|----------|-------------------------|
| 45 | Ex1 | Erichsen 07 E J Neurol |
| 46 | Ex5 | Erichsen 07 E J Neurol |
| 47 | Ex10_12 | Erichsen 07 E J Neurol |
| 48 | Ex2_17 | Erichsen 07 E J Neurol |
| 49 | Ex4_7 | Einsendung von Erichsen |
| 50 | Ex2_16 | Einsendung von Mannan |
| 51 | Ex2_7 | Einsendung von Klebe |
| 52 | Ex9_17 | Einsendung von Klebe |
| 53 | Ex4_9 | Einsendung von Schüle |
| 54 | Ex1 | Einsendung von Depienne |
| 55 | Ex1_17 | Einsendung von Depienne |
| 56 | Ex17 | Einsendung von Reid |
| 57 | Ex1_4 | Einsendung von Reid |
| 58 | Ex1 | Einsendung von Reid |
| 59 | Ex1 | Einsendung von Reid |
| 60 | Ex10_12 | Teil von REEP-Samplen |
| 61 | Ex2_7 | Teil von REEP-Samplen |
| 62 | Ex1 | Teil von REEP-Samplen |
| 63 | Ex1 | Teil von REEP-Samplen |

Tab. 1.1: Alle bekannten SPAST-Deletionen

1.4. Verantwortliche Mechanismen für große Deletionen

Im Unterschied zu Punktmutationen, welche sich vermutlich eher bei der DNA-Replikation oder Reparaturfehlern ereignen, resultieren größere Mutationen wie Deletionen, Duplikationen, Inversionen und Translokationen aus Rekombinationsmechanismen (Shaw und Lupski 2004). Dafür gibt es zwei unterschiedliche Vorstellungen: NAHR (*non-allelic homologous recombination*) und NHEJ (*non-homologous end joining* = Reparatur von mehrfachen Doppelstrangbrüchen).

1.4.1. Nicht allelische homologe Rekombination

Die nicht allelische homologe Rekombination (NAHR) ist ein Prozess, bei dem zwei DNA-Sequenzsubstrate, welche einen signifikanten Homologiebereich haben, sich zusammen lagern und es dadurch zu einem Strangaustausch oder -verlust kommt. Daraus resultieren Deletionen, Duplikationen und Translokationen (Lupski 2004). Homologiebereiche repräsentieren meist Repeats (Krawczak und Cooper 1991). Die Zusammenlagerung von Repeats, welche zu ungleichen Crossing-over-Ereignissen führt, scheint eine entscheidende Rolle für besonders große Gendeletionen zu spielen (Krawczak und Cooper 1991).

Deletionen-verursachende Repeats sind oft paraloge Sequenzen wie LCRs (*low copy repeats*) (Lupski 2004). Diese DNA-Abschnitte sind 10-400kb lang, besitzen eine 95-97%ige Homologie untereinander und machen 5% des Genoms aus (Stankiewicz und Lupski 2002b). Aber auch repetitive Elemente wie Alu-Sequenzen kommen in Frage, NAHR-Ereignisse zu vermitteln (Shaw und Lupski 2005, Deininger und Batzer 1999). Alus gehören zu den retrotransposablen Elementen. Sie sind ca. 300bp lang, kommen in mehr als 500000 Kopien weit verteilt im Genom vor und machen etwa 10% der Masse des menschlichen Genoms aus (Deininger und Batzer 1999). Es sind aber auch Assoziationen zwischen Alu-Elementen und LCRs beschrieben, bei denen sich Alu-Sequenzen im Bereich der Fusionssequenz bei LCR-Rekombinationen befinden (Shaw und Lupski 2004).

Um homologe Ereignisse zu vermitteln, ist eine starke Sequenzhomologie von 200-300bp zwischen zwei DNA-Substraten notwendig, was gerade der Größe der meisten Alu-Elemente entspricht. Die erforderliche Homologie wird als MEPS (*minimal efficient processing segment*) bezeichnet (Stankiewicz und Lupski 2002a).

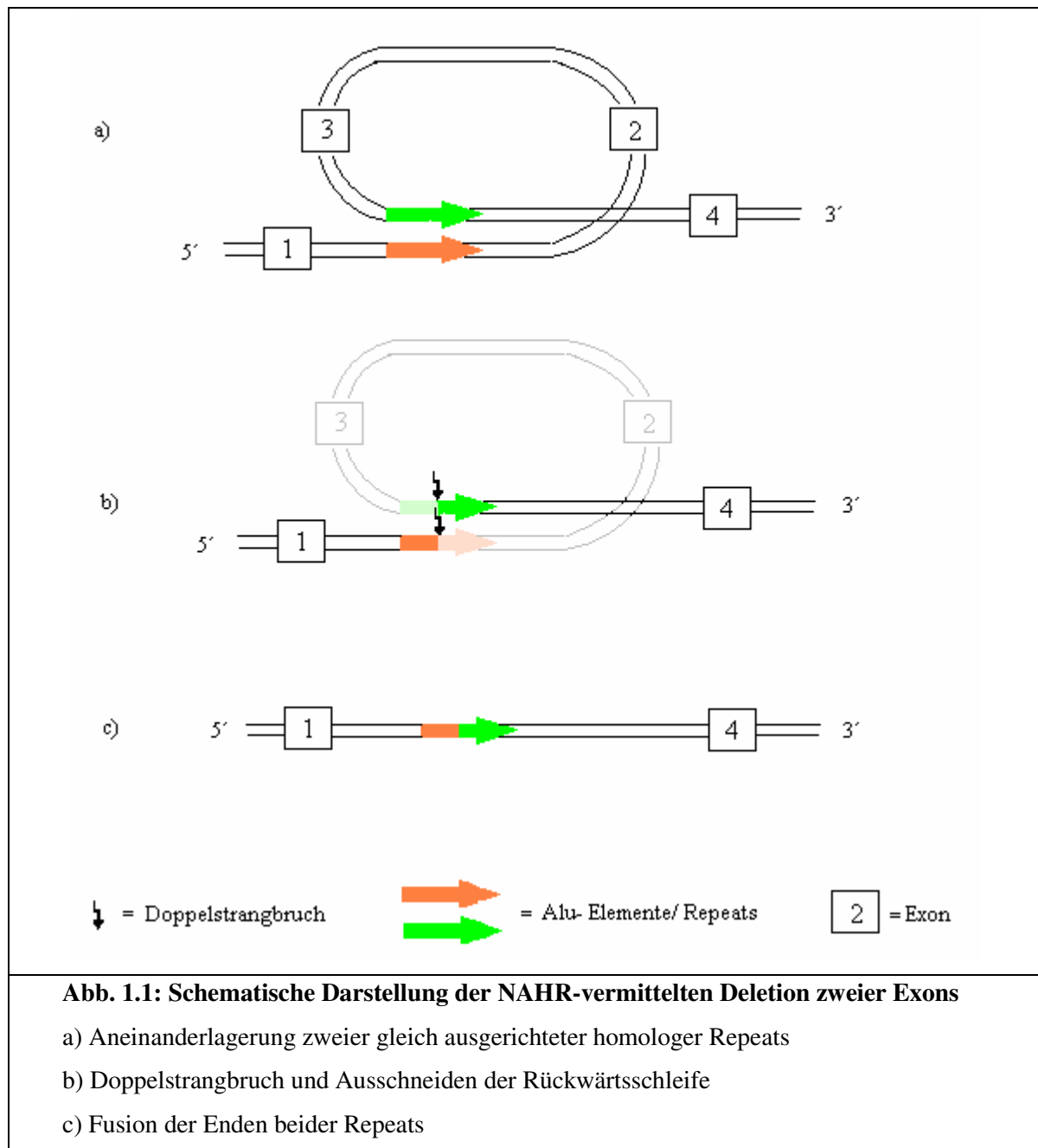
Durch eine bestimmte Ausrichtung von Chromosomen/Chromatiden können die nicht allelischen LCRs oder die repetitiven Elemente als Substrat für NAHR dienen. Die Wahrscheinlichkeit für ein Rekombinationsereignis scheint abhängig zu sein von dem Grad an Homologie, dem Abstand zwischen den Repeats und der Orientierung der Repeats zueinander (Stankiewicz und Lupski 2002a).

Der Strang austausch kann interchromosomal, intrachromosomal (zwischen den Schwesterchromatiden) oder innerhalb des Chromatids erfolgen. Er ereignet sich direkt, wenn die repetitiven Elemente die selbe Orientierung besitzen, und invertiert, wenn die Repeats entgegengesetzt ausgerichtet sind (Stankiewicz und Lupski 2002a). Während aus der interchromosomalen und intrachromosomalen Rekombination zwischen den Schwesterchromatiden jeweils duplizierte und deletierte Chromosomenabschnitte hervorgehen, führt die Ausbildung einer Rückwärtsschleife innerhalb eines Chromatids durch direkt zueinander orientierte Repeats zum Ausschneiden dieser Schleife und der daraus resultierenden Deletion (siehe Abb.1.1). Dagegen kommt es bei dem innerhalb eines Chromatids stattfindenden Rekombinationsereignis zwischen gegensätzlich ausgerichteten LCRs/repetitiven Elementen zu einer Inversion (Stankiewicz und Lupski 2002a).

Aktuell handelt es sich bei den meisten der bekannten genomischen Erkrankungen durch große DNA-Aberrationen um Deletionen (Stankiewicz und Lupski 2002a). Des Weiteren sind sowohl interchromosomal, intrachromosomal und innerhalb eines Chromatids komplexe

Rekombinationsereignisse möglich, welche zurückzuführen sind auf eine komplexe Struktur der beteiligten LCRs bzw. repetitiven Elemente (Stankiewicz und Lupski 2002a).

Alu-Alu-Rekombinationen ereignen sich meist zwischen Kopien auf demselben Chromosom und führen zu Deletionen oder Duplikationen. Ein Ereignis zwischen verschiedenen Chromosomen ist aber auch möglich mit dem Resultat von Translokationen oder komplexeren DNA-Veränderungen (Deininger und Batzer 1999).



Allgemein kann festgestellt werden, dass außer den Alu-Elementen auch andere repetitive DNA-Elemente zu genomischen Erkrankungen durch Retrotransposition und durch ungleiche homologe Rekombination zwischen den Elementen beitragen (Smit 1996). Dabei rekombinieren sich längere Repeats mit hoher Sequenzidentität leichter als kurze Repeats wie Alus, die zu einer fehleranfälligen Rekombination neigen (Deininger und Batzer 1999). Innerhalb der LCRs gibt es Regionen, in denen häufiger ein Strang austausch stattfindet, was einen Cluster an Rekombinations-Bruchpunkten liefert. Damit sind Rekombinations-Hotspots ein typisches Merkmal der nicht allelischen homologen Rekombination (Lupski 2004, Shaw und Lupski 2004). Selbst innerhalb der nur 300bp langen Alus wurden Hotspots beschrieben (Deininger und Batzer 1999).

1.4.2. Reparatur von mehrfachen Doppelstrangbrüchen

NHEJ (*non-homologous end joining*) ist ein echter Reparaturmechanismus zur Wiederherstellung der DNA-Integrität nach Doppelstrangbrüchen. Es beinhaltet den Doppelstrangbruch an zwei Stellen und das Wiederausfüllen der Lücken, wobei der DNA-Abschnitt zwischen den zwei Bruchpunkten verloren geht (Tsukamoto und Ikeda 1998, Roth et al. 1985). Abb. 1.2 stellt diesen Vorgang schematisch dar.

NHEJ-Bruchpunkte scheinen im Gegensatz zu NAHR-Bruchpunkten nicht in LCRs oder Alu-Elementen zu clustern, sondern befinden sich in nicht repetitiver Sequenz (Shaw und Lupski 2004). Die resultierenden Deletionen variieren in ihrer Größe und die Bruchpunkte sind weit verteilt und meist nicht wiederkehrend (Shaw und Lupski 2004). Trotzdem können auch LCRs oder Alu-Elemente in der Nähe der Bruchpunkte gefunden werden, aber nicht im Sinne einer Alu-/LCR-vermittelten Rekombination, sondern einer Prädisposition für DNA-Doppelstrangbrüche (Shaw und Lupski 2004).

Dabei können die beiden Bruchenden wieder verbunden werden, ohne dass eine ausgedehnte Homologie zwischen den Enden notwendig ist (Tsukamoto und Ikeda 1998). 40% der berichteten Verbindungen zeigen keine ausgedehnte Homologie (Roth et al. 1985). Typisch ist aber eine Mikrohomologie von 1-6nt am Fusionspunkt, wobei diese Übereinstimmung das Wiedervereinigen möglicherweise erleichtert (Roth und Wilson 1986, Roth et al. 1985).

Manchmal kann es erforderlich sein, dass die Bruchenden vor dem Zusammenfügen prozessiert werden müssen (Tsukamoto und Ikeda 1998). Dabei kann es zum Verlust oder zur Insertion einiger Nukleotide kommen. Damit ist das NHEJ ein fehleranfälliger

Reparaturmechanismus hinsichtlich der Integrität der DNA an den Fusionsstellen (Toffolatti et al. 2002, Pastink et al. 2001, Barnes 2001).

Auch zu NHEJ gibt es verschiedene Vorstellungen über den Ablauf, welcher abhängig ist von den jeweiligen strukturellen und sequenziellen Gegebenheiten (Tsukamoto und Ikeda 1998). Wenn die Enden als Einzelstränge vorliegen und sich dort komplementäre Sequenzen befinden, kommt es direkt zur Basenpaarung, gefolgt von einer vervollständigenden DNA-Synthese an den Einzelstrangstümpfen und anschließender DNA-Ligation der neu synthetisierten DNA an die Doppelstrangenden (Tsukamoto und Ikeda 1998). Wenn keine komplementären Sequenzen an den Einzelstrangüberhängen vorhanden sind, kann es auch zuerst zu einer Ligation der Einzelstränge kommen mit anschließender DNA-Synthese und Ligation der neu synthetisierten DNA-Enden. Hierbei ist keine Homologie erforderlich (Tsukamoto und Ikeda 1998). Ein neues Modell beinhaltet eine DNA-Synthese von einigen wenigen Nukleotiden am Einzelstrangüberhang eines der beiden Enden ohne das Vorliegen einer Matrize. Dies befähigt dieses Ende dann zur Basenpaarung mit dem anderen Ende, gefolgt von DNA-Synthese und Ligation. Auch hierfür sind komplementäre Nukleotide nicht die Voraussetzung (Tsukamoto und Ikeda 1998). Das Ergebnis ist dann eine Insertionsverbindung (Roth et al. 1985). Als Resultat sind nun Verbindungen mit limitierter Homologie, Verbindungen mit Insertionen oder das bloße Zusammenfügen ohne Homologie möglich (Krawczak und Cooper 1991). Insgesamt kann jedoch festgestellt werden, dass die meisten Rekombinationen sich zwischen DNA-Enden mit einigen wenigen Basen identischer Sequenz ereignen (Tsukamoto und Ikeda 1998). So wurde das Vorhandensein von 1-6nt Homologie bei 60% der analysierten nicht homologen Verbindungssequenzen gefunden (Roth und Wilson 1986). Dennoch ist es schwierig, aufgrund der nur sehr kurzen Homologie zu unterscheiden, ob die Homologie von mechanistischer Relevanz ist oder nur die Wahrscheinlichkeit des Homologieereignisses widerspiegelt (Roth et al. 1985). Es wurde aber durch eine statistische Analyse gezeigt, dass eine kurze Homologie häufiger vorkommt, als es die Wahrscheinlichkeit allein erklären würde (Roth und Wilson 1986).

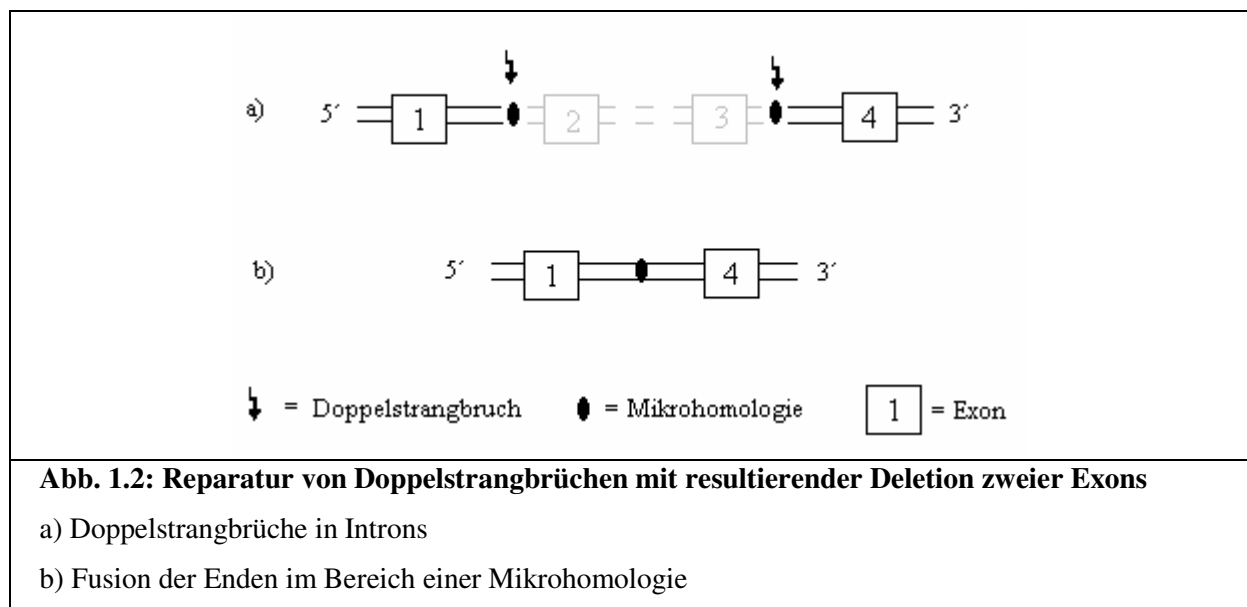
Wie schon erwähnt, repariert NHEJ ungewollte Doppelstrangbrüche, was nicht gleichbedeutend mit rein zufälligen Ereignissen ist. So wird die Rolle von Sequenzmotiven wie *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats* (Abeyasinghe et al. 2003) und von sekundären Strukturmerkmalen wie Haarnadelschleifen (Chuzhanova et al. 2003) bei der Generierung von Doppelstrangbrüchen diskutiert. So können z.B. Palindrome die Ausbildung von solchen Sekundärstrukturen beeinflussen, wodurch der effektive Abstand zwischen

kurzen Homologien verringert werden soll und damit der Strangbruch erleichtert werden kann (Krawczak und Cooper 1991).

Da bei Abeysinghe et al. 2003 bei einem beträchtlichen Anteil der Bruchpunktverbindungen von Deletionen invertierte Repeats (Palindrome), Inversionen von invertierten Repeats und symmetrische Elemente gefunden wurden, welche z.B. die Ausbildung von Haarnadelschleifen, Knoten-Strukturen und Möbius-Schleifen ermöglichen, wurde vorgeschlagen, dass diese Sekundärstrukturen mit großen Rekombinationsereignissen assoziiert sind (Chuzhanova et al. 2003).

Außerdem sollen Deletions-Bruchpunkte in AT-reicher Umgebung liegen und alternierende Purin-/Pyrimidinstrecken signifikant überrepräsentiert sein (Chuzhanova et al. 2003). Als rekombinogenische Motive fanden Abeysinghe et al. 2003 ein *Immunglobulin-Heavy-Chain-Class-Switch-Repeat* von fünf Nukleotiden in der 15bp-Umgebung der Deletions-Bruchpunkte als signifikant überrepräsentiert sowie auch Topoisomerase-Schnittstellen in der Nähe der Bruchpunkte. Weiter wurden *matrix attachment regions* (MARs), Homologe der Translin-Bindestelle oder die *chi-like*-Sequenz mit dem NHEJ in Zusammenhang gebracht, was aber nicht automatisch bedeutet, dass diese Motive direkt in den Rekombinationsprozess involviert sind (Toffolatti et al. 2002, McNaughton et al. 1998, Sironi et al. 2003, Matejas et al. 2006, Abeysinghe et al. 2003).

Somit scheint NHEJ nicht von repetitiven Elementen abhängig zu sein. Dennoch besteht eine Assoziation zu spezifischen Strukturen.



2. Zielstellung

Die vorliegende Arbeit hat eine detaillierte Charakterisierung bekannter Deletionen im SPAST-Gen zum Inhalt. Konkret bedeutet dies eine vergleichende Zusammenstellung der Verteilung aller Deletionsendigungen sowie die Amplifizierung und Sequenzierung ausgewählter Fusionen. Dies soll dazu dienen, die folgenden Fragen zu beantworten:

- (i) Gibt es Hinweise für eine Assoziation von Deletionsereignissen mit bestimmten SPAST-Regionen?
- (ii) Welche Mechanismen liegen solchen Ereignissen zu Grunde?
- (iii) Welche Ähnlichkeiten und welche Unterschiede bestehen zu anderen, häufig von Deletionen betroffenen Genen?

Die Beantwortung dieser Fragen ist sowohl für die molekulargenetische SPAST-Diagnostik als auch für ein besseres Verständnis des Mutationsmechanismus für Deletionen von Relevanz.

3. Material und Methoden

3.1. Geräte und Verbrauchsmaterialien

3.1.1. Chemikalien

| | |
|--|--|
| Agar Agar | FLUKA Chemie GmbH (Buchs, CH) |
| Ampicillin | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| APS (= Ammoniumpersulfat) | Appli Chem (Darmstadt, BRD) |
| Borsäure | Carl Roth GmbH (Karlsruhe, BRD) |
| Chill-Out™ 14 Liquid Wax | MJ Research™ Inc. (Boston, USA) |
| Destilliertes Wasser | B. Braun Melsungen AG (Melsungen, BRD) |
| DMSO (= Dimethylsulfoxid) | Sigma® Aldrich (Steinheim, BRD) |
| EDTA (= Ethylendiamin-Tetraacetat) | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| EDTA disodium salt | Sigma® Aldrich (Steinheim, BRD) |
| Essigsäure (100%) | Carl Roth GmbH (Karlsruhe, BRD) |
| Ethanol | Gibco (Eggenstein, BRD) |
| Ethidiumbromid | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| Hefeextrakt | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| Glycerol ($\geq 98\%$) | Carl Roth GmbH (Karlsruhe, BRD) |
| IPTG (= Isopropyl-Thio- β -D-Galactopyranosid (100mM)) | Invitrogen (Carlsbad, USA) |
| Kanamycin | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| 2-Propanol | Merck Schuchardt OHG (Hohenbrunn, BRD) |
| Saekem LE Agarose | Biozym Scientific GmbH (Oldendorf, BRD) |
| Sucrose | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| TEMED (= N'N'N'N'-Tetramethylethylenediamine, 99%) | Sigma® Aldrich (Steinheim, BRD) |
| Trisbase | Sigma® Aldrich (Steinheim, BRD) |
| Tris-HCl | Sigma® Aldrich (Steinheim, BRD) |
| Pepton | Serva Boehringer Ingelheim (Heidelberg, BRD) |
| Urea | Sigma® Aldrich (Steinheim, BRD) |
| x-Gal (= 2% Dimethylformamid) | Gibco BRL (New York, USA) |
| Xylene Cyano CF7 | Serva Boehringer Ingelheim (Heidelberg, BRD) |

3.1.2. Geräte

Elektrophoresekammer

Hoefer Super Sub

Amersham Biosciences Europe GmbH
(Freiburg, BRD)

Gefäße

Standard Eppendorf Tubes-3810X, 1,5ml

Eppendorf AG (Hamburg, BRD)

PP-test tubes-50ml (sterile) Celestar

Greiner bio-one (Kremsmünster, A)

Heizblock

FALC Thermoblock

Renner GmbH (Darmstadt, BRD)

PCR-Cycler

Mastercycler gradient

Eppendorf AG (Hamburg, BRD)

Hybais-Thermal Cycler

Thermo Life Sciences (Egelsbach, BRD)

PCR-Platten

Thermowell, 96 Well Plate

Corning Inc. Costar (Schiphol-Rijk, NE)

Pipetten

Lab Mate-Pipetten

ABIMED HAT (Virginia, USA)

Schüttelinkubator

Schüttelinkubator 3033

GFL mbH (Burgwedel, BRD)

Sequenzierer

Licor Long Reader 4200

Licor Bioscience (Nebraska, USA)

Spannungsgeräte

Elektrophoresis Power Supply EPS 301

Amersham Biosciences Europe GmbH
(Freiburg, BRD)

| | |
|--|---|
| Spektralphotometer GeneQuantil | Pharmacia Biotech (Cambridge, UK) |
| Tischzentrifuge | Eppendorf AG (Hamburg, BRD) |
| Transilluminator Image Master VDS | Pharmacia Biotech (Cambridge, UK) |
| Vortexer Vortex Genie 2 | Scientific Industries Inc. (New York, USA) |
| Wärmeschrank Wärmeschrank B6 | Heraeus instruments (Hanau, BRD) |
| Zentrifuge Eppendorf Centrifuge 5415R | Eppendorf AG (Hamburg, BRD) |

3.1.3. Verwendete Kits

| | |
|---|----------------------------------|
| Klonierungs-Kit Topo TA Cloning® for Sequencing Kit (Part no. 45-0030) | Invitrogen GmbH (Karlsruhe, BRD) |
| Gelaufreinigungs-Kit MinElute Gel Extraction (Kat.-No. 28606) | Qiagen GmbH (Hilden, BRD) |
| Plasmidaufreinigungs-Kit QIAprep Spin Miniprep Kit (250) (Kat.-No. 27106) | Qiagen GmbH (Hilden, BRD) |

Sequenzierungs-Kit

Thermo Sequenase Fluorescent Labelled

Primer Cyclor Sequencing-Kit

RPN 2438

Amersham Biosciences Europe GmbH

(Freiburg, BRD)

3.2. Patienten-DNA

Untersucht wurde die DNA von zehn nicht verwandten HSP-Patienten, für die das Vorliegen großer SPAST-Deletionen bekannt war (Depienne et al 2007). Die DNA-Proben wurden im November 2006 auf Anfrage von Dr. C. Depienne (INSERM U 679 Hôpital de la Salpêtrière, 75651 PARIS CEDEX 13, FR) zur Verfügung gestellt. Diese und weitere große SPAST-Deletionen wurden in Depienne et al, 2007, J Med Genet 44: 281-284, beschrieben. Es handelt sich dabei um folgende Aberrationen, die noch einmal mittels MLPA nachgewiesen wurden und in Tabelle 3.1 aufgeführt sind.

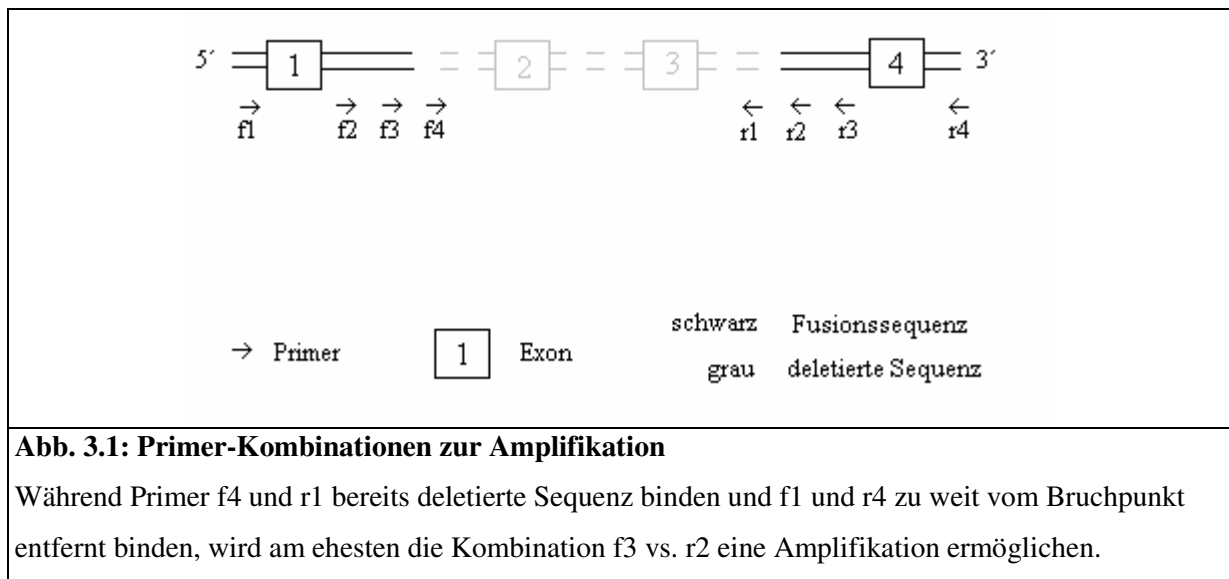
| Laufende Nr. | Bezeichnung | Deletion |
|--------------|-------------|------------|
| P1 | FSP 742 | delEx5_6 |
| P2 | FSP 530 | delEx5_7 |
| P3 | FSP 167 | delEx6 |
| P4 | FSP 374 | delEx8_12 |
| P5 | FSP 157 | delEx9 |
| P6 | FSP 235 | delEx9_12 |
| P7 | FSP 335 | delEx10_16 |
| P8 | FSP 333 | delEx13 |
| P9 | FSP 236 | delEx16 |
| P10 | FSP 803 | delEx16 |

Tab. 3.1: Aberrationen der untersuchten DNA-Proben

Die erhaltenen Proben enthielten bereits aufbereitete DNA-Lösungen, welche durch eine Standard-Aussalzungsmethodik aus Leukozyten (EDTA-Blut) gewonnen worden waren. Es wurden jeweils 15µl DNA-Lösung mit einer Konzentration von 200ng/µl erhalten, die für die PCR-Amplifikationen 1:20 (10µl auf 190µl destilliertes Wasser) auf 10ng/µl verdünnt wurden.

3.3. Primerdesign und Primer

Die Primersequenzen wurden so gewählt, dass sie Sequenzen in benachbarten Introns zu dem jeweils ersten und letzten deletierten Exon des jeweiligen Templates binden. Es wurden jeweils mehrere Primersequenzen pro Intron ausgewählt, die einen Abstand von 1-2kb haben (Abb. 3.1).



Um unspezifische Amplifikationen zu vermeiden, sollten die Bindestellen nicht innerhalb repetitiver Sequenzen liegen. Dafür wurden die beteiligten Introns auf repetitive Sequenzelemente mit dem Programm *Repeatmasker* (<http://www.repeatmasker.org/>) untersucht.

Ein weiteres Kriterium war das Erreichen eines möglichst ausgewogenen Verhältnisses aus A/T zu G/C ($\approx 1:1$). Außerdem sollte das 5'-Ende der Primer G oder C sein, da hier die Bindung zum Template besonders stark ist wegen der Ausbildung von drei Wasserstoffbrückenbindungen bei einer G-C-Bindung im Vergleich zu nur zwei Wasserstoffbrückenbindungen zwischen A-T-Bindungen.

Die Länge der Primer wurde nach der „4+2-Regel“ bestimmt. Ziel war eine Hybridisierungstemperatur von 58-60°C. Diese erhält man durch Addition von 4K für jedes enthaltene G oder C und 2K für A oder T. Demzufolge lag die Länge der letztlich gewählten Primersequenzen zwischen 20-28 Nukleotiden (Tabelle 3.2).

| Primer-Bezeichnung | Sequenz 5' - 3' |
|--------------------|-------------------------------------|
| Int04_f01 | AGA AGA TGC AAC CAG TTT TGC |
| Int04_f02 | TCT CTG ACA TCT GAG CTA GC |
| Int04_f03 | CCA GAA CAA CCT TAG GTT AGG |
| Int04_f04 | CTC TTC CTG CAA CCT GTT TG |
| Int04_f05 | GTT TCA TAT ACA ATG CTT TAA CTC |
| Int04_f06 | GAT AAT CAT TGA CTC CTT ATG G |
| Int04_f07 | GTA TAT ATA CAG TAT ACT ATA TCG TG |
| Int04_f08 | ATT CTC ATA CTT GTT CCT TGT G |
| Int04_f09 | TAC AGC CCA TAT AAT ACT TAG C |
| Int04_f10 | CCT ATC TAC CTA GTG ACC AC |
| Int05_f01 | CTA CAA TTT TCT AAT CAC AAT GG |
| Int05_f02 | AAG TTT TGC TTA CCC AGT CTG |
| Int07_f01 | TCA TAG GGC TTA GGC TTC ATC |
| Int07_f02 | ATC GAA GGC TAA ATA CCA ACG |
| Int07_f03 | CCT GGT TAG TGC TAT AGC TG |
| Int07_f04 | TTT ATA GAA GAA TGT GAT TGT TCT C |
| Int07_f05 | TGA TAT TTT ATA ATG ATA TGC TTG ATG |
| Int07_f06 | GAC TTT GTT TTC TGT TAG GTT GAG |
| Int07_f07 | CAA CAA ATG GTA TAG GAC AAC TG |
| Int07_f08 | TAT GGG AAC TCT GTA CTT TCT G |
| Int07_f09 | AGT GAT TTA CCA CAC TTT ATC AG |
| Int07_f10 | AAA GAG TAC TTA AAA TGT CTC TAG |
| Int09_f01 | TTA AAT CGG TAA ATA TGG TTA TC |
| Int09_f02 | GAT TCT GAT CTA AAA TAT GAC AG |
| Int09_f03 | TTG CCT TTA TCA AAA CCT TTA TG |
| Int09_f04 | GCA CCG AAA TCC AGA GGT G |
| Int09_f05 | TTT GAA GTC TGT TTG CTA GAG C |
| Int09_f06 | TGC CTG AAT ATC TTG AAG TAA G |
| Int09_f07 | TTA CAT ATG TAT CTT AAC ATG TTG |
| Int09_f08 | TGA AAT TTT ACA AGT CGT TAC TC |
| Int09_f09 | AAG TTG TTA TTA AAG CAT GTT TAC |
| Int09_f10 | ATT GAC CTT CTG GAC TAT CTG |
| Int09_f11 | TCT CAA GTC TTA GGA ATG CAG |
| Int12_f01 | TTA AAT TCC TGT GTG CTA GAT TTT C |
| Int12_f02 | CAC AAC CTC ACC AAC ATC TG |
| Int12_f03 | GAG TCA AAT CAA GAA CAC AGT C |
| Int12_f04 | ATT TCA CTG GGC TCT TGT TTC |
| Int12_f05 | AAT TCT AAC AAT ACC GAA TTA AGG |
| Int15_f01 | GAG GGG AAA TAA TTT GCT GTT TC |
| Int15_f02 | TTT TGG CTA TTC TAA GGT AGA C |
| Int15_f03 | ATA CAT GTA GAT CAT TGT ACT TG |
| Int05_r01 | CAA AAC TCT AGG TAT GTT ACT G |
| Int05_r02 | ATT TAG ATG GCA AAA CTT ACT TG |
| Int07_r01 | AAA TCT CTC TCA CCT TGA TAT C |
| Int07_r02 | TAT CTA CTT TGA TGA CTT AAT ATC |
| Int07_r03 | GGC AGC TCT AAT CTT AGA AAG |
| Int07_r04 | TTA AGC TTA TAT GAA AGA TAA CAT C |
| Int07_r05 | GGA TGA ATA TCT AGA AAG GTC C |
| Int07_r06 | TTG ACC TAT AAA AAC AGT AAT TTC AG |
| Int07_r07 | ATA ACA GTG AAG GTG TAT GTA C |
| Int07_r08 | AAA CAA GAC ATA TTG AAC ATA GG |
| Int07_r09 | CCC AAA TAT CTC AAC AGA AAA C |

| Primer-Bezeichnung | Sequenz 5' - 3' |
|--------------------|--|
| Int09_r01 | TCA TAT GTA AAT AAT AGA TAC AAT AG |
| Int09_r02 | TAT TTA CAC ACT AGT TCT GCA C |
| Int09_r03 | TGG TGA TCA CCT CTG GAT TC |
| Int09_r04 | ACA TGA AGA TGA AAC AGT ACT C |
| Int09_r05 | TTG TCC ATG ACT ATG ATG CAG |
| Int09_r06 | GTA AAC ATG CTT TAA TAA CAA CTT G |
| Int09_r07 | AAA AGA AAT GTC ACC AGA TTT AG |
| Int09_r08 | ATT TCA CAA ATA GTC AAT AAC TTA G |
| Int12_r01 | GTA ACC CAT CCA GTA AAG AG |
| Int12_r02 | TGC ACA GCA AAA AGA ACC ATC |
| Int12_r03 | GCA ACA TAA CAC ATA ATT TAT GTG |
| Int12_r04 | AAT GAA GAA CAT GCA TGA GAA G |
| Int12_r05 | ATA TTC AAG GCA CTG GAA TGC |
| Int12_r06 | GAG GTA TGA ATT CAT CTA CTA TG |
| Int12_r07 | ATC TGT GTT AGT GTG CAA TAT G |
| Int12_r08 | CTA AGG TGG TAT AAT AAA TAG TTC |
| Int13_r01 | ATG GTA GTT CTT GTT TCT GCT C |
| Int13_r02 | CAC TGA TCT GCT TTA CGT TTC |
| Int13_r03 | ATC AAT GAA TTA CTC ATT TCA TAC |
| Intr16_r01 | GTC TTA TAT ATG TCT TAC AAT ATA GAA G |
| Intr16_r02 | GTT AAT TTT TTC TAG CTT CTA ATC AAC |
| Intr16_r03 | TAG TTA TAG GTT ATT TTT ATT GTA TAA AC |
| Intr16_r04 | TTA TCT GTG TTA ACA GTT CAC TG |
| Intr16_r05 | ATT CTT CCT TGT AAA TCA CTC AG |
| Intr16_r06 | ATT TGA ACT TCT ACT ATT ATA AAC TG |
| Intr16_r07 | TCA TCT TTG ATA TCT CCA TGA G |
| Intr16_r08 | TTA AAC AGT GGT ATC TCC AAA GTC |

Tab. 3.2: Alle verwendeten Primer

3.4. PCR-Screening

Material:

1. genomische DNA-Verdünnungen (10ng/μl)
2. 10x Puffer Bio Therm (15m MgCl₂) von GeneCraft GmbH (Lüdinghausen, BRD)
3. dNTP-Gemisch (je 10mM) von GeneCraft GmbH (Lüdinghausen, BRD)
4. Polymerase TaKaRa LA Taq TM von TaKaRa Bio. Inc. (Otsu, Shiga, Japan)
5. Primer (forward und reverse) (10pmol/μl) von MWG Biotech AG
(Ebersberg, BRD)
6. PCR-Cycler

Ziel der Durchführung der PCRs war die Amplifikation der Deletions-Fusionssequenzen, die die Bruchpunkte enthalten, sowie das Herausfinden derjenigen Primer-Kombinationen, die eine Amplifikation erst ermöglichen. Eine Amplifikation war nur dann möglich, wenn forward- und reverse-Primer sich nahe genug zueinander anlagerten und andererseits nicht in bereits deletierter Sequenz lokalisiert waren (Abb. 3.1).

Folgender PCR-Ansatz wurde für eine Primer-Kombination gewählt:

| Substanz | Menge in μl |
|----------------|------------------------|
| Wasser | 13,30 |
| 10x Puffer | 2,00 |
| dNTPs | 0,50 |
| DNA | 1,00 |
| Taq-Polymerase | 0,20 |
| P _f | 1,00 |
| P _r | 1,00 |
| Σ | 20,00 |

Tab. 3.3: PCR-Ansatz

Der Ansatz wurde gemischt und auf PCR-Platten pipettiert. Anschließend wurden 20 μl Wachs auf den PCR-Ansatz gegeben.

Wenn in einem Ansatz mehrere Primer-Kombinationen für ein Template getestet wurden, vervielfachte sich die Menge von DNA, Puffer, dNTPs und Polymerase in einem Supermix. Die Substanzen wurden dann ohne Primer-Zugabe gemischt, jeweils ein Teil davon auf die PCR-Platte pipettiert und erst dann die Primer-Kombinationen in das jeweilige Gefäß auf der PCR-Platte gegeben.

Mit folgendem PCR-Programm wurde amplifiziert:

| Wiederholungen | Schritt | Temperatur in $^{\circ}\text{C}$ | Zeit |
|----------------|-------------------------------|----------------------------------|------|
| 1 | Initiale Denaturierung | 95 | 5' |
| 34 | Denaturierung | 95 | 1' |
| | Hybridisierung | 57 | 1' |
| | Elongation | 72 | 2' |
| 1 | Finale Elongation | 72 | 5' |

Tab. 3.4: PCR-Programm

Ziel der finalen Elongation war es, alle entstandenen Produkte mit einem A-Überhang zu versehen, um das Klonieren in den TOPO-TA-Vektor (siehe 3.5.3.1.) zu erleichtern.

3.5. Charakterisierung der PCR-Produkte

3.5.1. Agarosegelelektrophorese

Material:

1. Agarose
2. 50x TAE-Puffer: Essigsäure (100%) 57,10ml
Tris 242,00g
EDTA 14,70g
3. Gelladepuffer: Sucrose 4,00g
2M Tris-HCl 0,50ml
0,5M EDTA 2,00ml
Xylene Cyano CF7 4,00mg
mit destilliertem Wasser auf 10,00ml aufgefüllt
4. DNA-Marker: 1kb-DNA-Ladder II (1µg/10µl) von GeneCraft GmbH
(Lüdinghausen, BRD)
5. Elektrophoresekammer
6. Spannungsgerät
7. Transilluminator

Mit der Agarosegelelektrophorese wurden die PCR-Produkte aufgetrennt, um sie dann zu visualisieren und weiterbearbeiten zu können.

Zunächst wurde ein Gel gegossen. Dabei wurde der 50x TAE-Puffer 1:50 (100ml 50x TAE zu 4900ml destilliertem Wasser) verdünnt und man erhielt 1x TAE-Puffer.

Zu 400ml 1x TAE-Puffer wurden für ein 1%iges Gel 4g Agarose hinzugegeben. Als Standard wurden 1%ige Gele verwendet. Wurden aber Fragmente <500bp erwartet, wurde ein 2%iges Gel hergestellt, indem die Menge der Agarose verdoppelt wurde.

Zum Auflösen der Agarose wurde das Gemisch in einem Mikrowellengerät zum Kochen gebracht. Anschließend wurden 8µl Ethidiumbromid dazugegeben und gemischt. Das Gemisch wurde in eine Gelkammer mit eingehängtem Kamm gegossen. Nach 30 Minuten Auspolymerisierung wurde der Kamm gezogen und das Gel mit 1x TAE-Puffer übergossen.

Zu den PCR-Produkten auf der PCR-Platte wurden jeweils 4µl Gelladepuffer pipettiert und die kompletten Proben (ca. 24µl) in jeweils eine Geltasche gefüllt. In eine Tasche wurden 8µl eines 1kb-Markers als Standardmolekulargewicht hinein gegeben.

Die Gelkammer wurde verschlossen und eine Spannung von 130V angelegt sowie eine Stromstärke von 400mA eingestellt. Nach 30-45 Minuten wurde das Gel unter UV-Bestrahlung in einem Transilluminator visualisiert.

3.5.2. Gelaufreinigung

Wenn Erfolg versprechende Banden dargestellt werden konnten, wurden diese aus dem Gel ausgeschnitten und die Gelstücke mit der darin enthaltenen DNA aufgereinigt, um diese weiterbearbeiten zu können. Die Gelaufreinigung erfolgte mit dem Kit „MinElute Gel Extraction“ entsprechend der Herstellerangaben von Qiagen:

- nach dem Ausschneiden Gelstücke gewogen und in Eppendorf-Tubes gegeben
- drei Volumeneinheiten vom Puffer QG zu einer Gewichtseinheit der Gelstücke (z.B. 100µl auf 100mg) gegeben
- bei 50°C zehn Minuten im Heizblock inkubiert, dabei gemischt
- eine Volumeneinheit Isopropanol dazugegeben und gemischt
- mitgelieferte Säulen „MinElute spin column“ in weitere Eppendorf-Tubes gesetzt und Gemisch auf die Säulen pipettiert, eine Minute in Tischzentrifuge bei 13200x g zentrifugiert und anschließend die Durchläufe verworfen
- 0,5ml Puffer QG auf die Säulen gegeben und eine Minute zentrifugiert, Durchläufe verworfen
- 0,75ml Puffer PE auf die Säulen gegeben und eine Minute zentrifugiert, Durchläufe verworfen
- drei Minuten gewartet
- eine Minute zentrifugiert
- Säulen in neue Eppendorf-Tubes gesetzt und zum Eluieren 10µl Puffer EB auf die Membranen der Säulen pipettiert, eine Minute gewartet und dann eine Minute zentrifugiert.

3.5.3. Klonierung

Die Klonierung der aufgereinigten DNA-Fragmente erfolgte, um die Sequenzierung zu erleichtern. Dafür verfügte der verwendete Vektor über eine universale Sequenz im Bereich der Ligationsstelle mit einem DNA-Fragment. Für die Sequenzierung konnten dann universelle Primer benutzt werden.

3.5.3.1. Topoisomerasereaktion und Transformation

Material:

1. Topo TA Cloning® for Sequencing Kit von Invitrogen GmbH (Karlsruhe, BRD)
(Kat. No. K4575-01) bestehend aus: pCR®4-Topo (Vektor)

Salzlösung (1,2M NaCl; 0,06M MgCl₂)

S.O.C. Medium

One Shot® Top chemically competent

Escherichia coli

2. aufgereinigte DNA-Fragmente

3. Agarplatten aus: LB-Medium (1% Pepton; 0,5% Hefe; 0,5% NaCl; pH=7,2)

15,0g Agar Agar pro Liter LB-Medium

2,0ml x-Gal (0,02g/ml Dimethylformamid)

2,0ml IPTG (23,8mg/ml destilliertes Wasser)

1,0ml Kanamycin

4. Heizblock

5. Schüttelinkubator

6. Wärmeschränk

Für die Topoisomerasereaktion wurde folgender Ansatz gemischt:

| Substanz | Menge in µl |
|------------|-------------|
| Wasser | 2,00 |
| Vektor | 0,33 |
| Salzlösung | 0,66 |
| DNA-Lösung | 1,00 |
| Σ | 4,00 |

Tab. 3.5: Ansatz für Topoisomerasereaktion

Dann wurden zur Transformation 25µl E.coli-Bakterien langsam auf Eis aufgetaut. Der 4µl-Ansatz der Topoisomerasereaktion wurde zu den Bakterien gegeben und 20 Minuten auf Eis inkubiert. Dann erfolgte eine Minute Inkubation bei 42°C im Heizblock, bevor weitere zwei Minuten auf Eis inkubiert wurde. 200µl S.O.C. Medium wurden dazu gegeben und dieser Ansatz wurde für eine Stunde in den Schüttelinkubator bei 37°C gestellt. Anschließend wurde der Bakterienansatz auf die Agarplatten gegeben, verteilt und über Nacht in den Wärmeschränk (37°C) gestellt.

3.5.3.2. Kolonie-PCR

Die Materialien entsprachen denen von 3.4. mit der Ausnahme, dass andere (universale) Primer verwendet wurden und dass das Template nicht in Form einer DNA-Lösung vorlag, sondern dass mit (sterilen) Holzspateln Proben von weißen Kolonien entnommen wurden. Weiße Kolonien waren ein Hinweis darauf, dass ein DNA-Fragment in den Vektor eingebracht wurde, welches den Leserahmen des LacZ α -Gens des Vektors unterbricht. Damit fehlte die β -Galaktosidase-Aktivität und x-Gal im Agar konnte nicht gespalten werden, wodurch eine blaue Farbreaktion unterblieb und die entsprechende Bakterienkolonie sich weiß darstellte. Da es sich bei den eingefügten Fragmenten aber nicht zwangsläufig um die gewünschten Fragmente handelte, wurden noch Kolonie-PCRs der weißen Kolonien zur Validierung durchgeführt. Die Größen der Amplikons sollten dabei denen, zuzüglich zur Länge der universalen Sequenz, von 3.4. entsprechen. Somit reduzierten die Kolonie-PCRs die Anzahl aufzuarbeitender unerwünschter Klone.

Folgender Ansatz wurde verwendet:

| Substanz | Menge in µl |
|----------------|-------------|
| Wasser | 5,0 |
| 10x Puffer | 0,8 |
| dNTPs | 0,5 |
| M13f-Primer | 0,8 |
| M13r-Primer | 0,8 |
| Taq-Polymerase | 0,1 |
| Σ | 8,0 |

Tab. 3.6: Kolonie-PCR-Ansatz

Dieser wurde gemischt, auf eine PCR-Platte pipettiert und ein Abstrich von den Kolonien wurde anschließend dazugegeben. Jeweils 10µl Wachs wurden auf die Ansätze gegeben.

Das PCR-Programm war identisch zu dem von 3.4. Anschließend erfolgte die Auftrennung von Fragmenten und Visualisierung wie bei 3.5.1. mit der Ausnahme, dass aufgrund der Ansatzmenge nur 10µl der PCR-Produkte zur Agarosegelelektrophorese aufgetragen wurden.

3.5.3.3. Inkubation der Bakterien als Flüssigkultur

Material:

1. LB-Medium (siehe 3.5.3.1)
2. Schüttelinkubator
3. Kanamycin
4. Reaktionsgefäße

Die im vorigen Schritt identifizierten Klone, die mit hoher Wahrscheinlichkeit die gesuchten DNA-Fragmente enthielten, wurden in Flüssigkulturen zur Vervielfachung der Plasmid-DNA vermehrt.

3ml LB-Medium und 150µl Kanamycin wurden in 50ml Gefäße gegeben. Mit sterilen Holzspateln wurden Bakterienkolonien dazugegeben. Dieser Ansatz wurde über Nacht bei 37°C in den Schüttelinkubator gestellt.

3.5.3.4. Plasmidaufreinigung

Material:

1. Glycerol (50%)
2. Zentrifuge
3. QIAprep Spin Miniprep Kit von Qiagen
4. Spektralphotometer

Die Plasmidaufreinigung war notwendig, um Sequenzen aus den Plasmiden sequenzieren zu können. Doch vor der eigentlichen Plasmidaufreinigung wurde jeweils ein Glycerolstock der Bakterienkolonien angelegt. Dies war ein Sicherheitsschritt, falls man beabsichtigte, einige der Kolonien noch einmal anzuziehen. 300µl jeder der Bakterienlösungen wurde mit 300µl Glycerol gemischt und bei -80°C eingefroren.

Die verbleibenden Bakterienkulturlösungen wurden 15 Minuten bei 3500x g zentrifugiert und die Überstände wurden abpipettiert und verworfen.

Es folgte die Plasmidaufreinigung entsprechend der Herstellerangaben:

- 250µl vom Puffer P1 zu den Bakterienzellen gegeben, gemischt und in mitgelieferte Mikrozentrifugen-Tubes gegeben
- 250µl Puffer P2 zugegeben und gemischt, bis Lösungen klar wurden
- 350µl Puffer N3 dazugegeben und gemischt
- zehn Minuten in der Tischzentrifuge bei 13200x g zentrifugiert
- die Überstände in Säulen (QIAprep spin column) pipettiert
- 30sec zentrifugiert und jeweils Durchlauf verworfen
- 0,5ml Puffer PB zugegeben, 30sec zentrifugiert und Durchläufe verworfen
- 0,75ml Puffer PE zugegeben, 30sec zentrifugiert und Durchläufe verworfen
- eine Minute zentrifugiert
- zum Eluieren Säulen in neue 1,5ml Mikrozentrifugen-Tubes eingehängt, 50µl Puffer EB auf die Mitte der Säulen gegeben, eine Minute gewartet und eine weitere Minute zentrifugiert

Anschließend wurde die eluierte Plasmid-DNA jeweils 1:50 verdünnt (2µl der DNA-Lösung zu 98µl destilliertem Wasser). Es folgte die photometrische Konzentrationsbestimmung der verdünnten DNA. Da alle Werte zwischen 100-500µg/ml lagen, konnte die sich anschließende Sequenzierung standardisiert mit den durch das Eluieren erhaltenen Konzentrationen durchgeführt werden.

3.5.4. Sequenzierung

Material:

1. Thermo Sequenase Fluorescent Labelled Primer Cycle Sequencing Kit
with 7-Deaza-dGTP von Amersham Biosciences Europe GmbH (Freiburg, BRD) (Kat.
Nr. RPN 2438) bestehend aus:
 - A-Reagenz
 - C-Reagenz
 - G-Reagenz
 - T-Reagenz
 - Loading dye
2. DMSO
3. Sequenzierprimer M13_universal oder M13_reverse (100pmol) von MWG Biotech
AG (Ebersberg, BRD) verdünnt auf 2pmol/μl
4. PCR-Cycler
5. 7%iger Long Ranger Gel-Ansatz:
 - Urea 12,60g
 - Long Ranger Gellösung 4,20ml
 - BMA (Rockland, USA)
 - auf 30 ml mit destilliertem Wasser aufgefüllt
6. APS (10%)
7. TEMED
8. 10x TBE-Puffer:
 - Trisbase 108,0g
 - Borsäure 55,0g
 - EDTA disodium salt 10,0g
 - auf 1000ml mit destilliertem Wasser aufgefüllt
9. 1x TBE-Puffer:
 - 10x TBE-Puffer 100ml
 - destilliertes Wasser 900ml
10. Ethanol
11. Sequenzierer

Sequenziert wurden definierte Bereiche aus den Plasmiden, welche die DNA-Fragmente aus den PCR-Amplifikationen enthielten. Ziel war die Basensequenzanalyse der Fusionsregionen. Damit war dieser Schritt die Voraussetzung für die Bruchpunkt-Analyse und weitere in silico-

Untersuchungen. Außerdem stellte die Sequenzierung eine endgültige Validierung der PCR-Amplifikation von den spezifischen SPAST-Deletions-Fusionsregionen dar.

Die A-, C-, G-, und T-Reagenzien wurden auf Eis aufgetaut. Es wurde ein Voransatz („5er Mix“) zusammengestellt aus: - destilliertem Wasser 10,00µl

- DMSO 1,25µl
- Sequenzierprimer 1,25µl

Dieser Ansatz wurde gemischt und jeweils 2,5µl dieses Ansatzes wurden zu jeweils 2,5µl A-, C-, G- und T-Reagenz gegeben. Auch diese Ansätze wurden gemischt (durch Auf- und Abpipettieren). Aus jedem dieser vier Ansätze wurde 1µl auf eine PCR-Platte pipettiert und jeweils 1µl DNA-Lösung dazugegeben. Dabei erfolgte das gesamte Zusammenstellen auf Eis. Anschließend wurden noch 10µl Wachs auf die Ansätze gegeben.

Die Reaktion erfolgte bei folgendem Programm:

| Wiederholungen | Schritt | Temperatur in °C | Zeit |
|----------------|------------------------|------------------|------|
| 1 | Initiale Denaturierung | 94 | 3′ |
| 35 | Denaturierung | 94 | 30″ |
| | Hybridisierung | 56 | 30″ |
| | Elongation | 72 | 45″ |
| 1 | Finale Elongation | 72 | 3′ |

Tab. 3.7: Programm der Sequenzierungsreaktion

Danach wurde zu den Produkten jeweils 2µl Loading dye gegeben und die Gemische wurden drei Minuten bei 95°C denaturiert und anschließend wieder auf Eis gebracht.

Es folgte die Anfertigung eines Acrylamidgels aus: - Long Ranger (7%) 20ml

- APS (10%) 140µl
- TEMED 14µl

Dieser Ansatz wurde gemischt und, nachdem die Glasplatten der Gel-Vorrichtung mit Ethanol gereinigt worden waren, wurde das Gemisch zwischen die beiden Glasplatten gegossen und ein Vorkamm eingehängt. Das Gel musste eine Stunde auspolymerisieren. Anschließend wurde der Vorkamm entfernt, dieser Bereich mit destilliertem Wasser umspült und der eigentliche Kamm eingeführt. Das Gel in den Glasplatten wurde in den Sequenzierer eingehängt und an den Enden mit 1x TBE-Puffer umspült. Es folgte die Einstellung der Spannung und des Lasers entsprechend der Herstellerangaben am angeschlossenen Rechner. Dann wurden jeweils 1,2µl der vier Proben in die Geltaschen pipettiert und der Sequenzierer gestartet. Auf dem Rechner wurden dann die einzelnen Banden entsprechend der einzelnen Fragmente visualisiert, woraus man die Sequenz ablesen konnte.

3.6. „Extra-Primer“ zur Validierung

Für einige Templates (P7, P8, P10), deren Sequenzen im Bereich der Fusion schlecht abzulesen waren, wurden weitere Primer bestellt. Je weiter die gesuchten Fusionsstellen von den M13-Sequenzen im Plasmid entfernt lagen, desto schwieriger war das Ablesen, da mit der Länge die Qualität der Sequenzierung abnahm. Diese neuen Primer besaßen teilweise (mindestens einer der jeweiligen Kombination) zusätzlich zu der das Template bindenden Sequenz noch einen universalen M13-Sequenz-Schwanz. Dieser ermöglichte eine Sequenzierung mit einem universalen M13-Primer ohne zwischengeschaltete Klonierung. Da diese Primer aber teuer sind, erfolgte dies nicht bei den initialen PCR-Amplifikationen, bei denen viele Primer-Kombinationen erst ausprobiert wurden um Banden zu erhalten und auch nicht bei allen Templates. Die PCR erfolgte wie bei 3.4. mit sich anschließender Agarosegelelektrophorese, Gelaufreinigung und Sequenzierung, wie unter 3.5. beschrieben. Folgende „Extra-Primer“ wurden zusätzlich verwendet:

| Primer-Bezeichnung | Sequenz 5' - 3' |
|--------------------|---|
| P7_bb_f | GAA AAG AAA AAG AAA ATT GAA ATG TCT AG |
| P7_bb_M13r_r | ACT GAC ACA GGA AAC AGC TAT GAC CCT GGG ATT ACA GGC GTG AG |
| P8_bb_f | CAA ATT GAT TTT TTC AAA TTT GTA AAA TAA C |
| P8_bb_M13r_r | ACT GAC ACA GGA AAC AGC TAT GAC CGC TGT ACC ATG GAT TGG AAG |
| P10_bb_M13f_f | ACT GAC TGT AAA ACG ACG GCC AGT GAG GTG AGG TTG CAG TAA G |
| P10_bb_M13r_r | ACT GAC ACA GGA AAC AGC TAT GAC CCC AAG GTC AGG AGA TCG AG |

Tab. 3.8: Alle verwendeten „Extra-Primer“

3.7. In silico-Analyse

3.7.1. BLAST

Nach dem Ablesen der Sequenzen wurden diese „geblastet“, d.h. in das BLAST-Programm folgender Internetadresse eingegeben: <http://genome.ucsc.edu>. Dieses Programm findet die genomische Lokalisation der eingegebenen Sequenz, liefert dessen Referenzsequenz und ermöglicht damit den Vergleich zwischen abgelesener und Referenzsequenz. Im Idealfall wurden zwei genomische Lokalisationen gefunden, die durch einen Abschnitt an Referenzsequenz, welche nicht in der abgelesenen Basenfolge vorkam, voneinander getrennt waren. Dieser Abschnitt entsprach den Deletionen und die Übergänge von vorhandener und nicht vorhandener Sequenz stellten die Bruchpunktregionen dar.

3.7.2. Homologie-Analyse

Die Übereinstimmung in den Sequenzumgebungen (50bp in 5'- und 3'-Richtung um die Bruchpunktregionen zusätzlich zur Bruchpunktregion) wurde zunächst mit dem Programm von <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi> untersucht. Dieses findet ausgeprägte Homologien zwischen zwei Sequenzen und gibt den prozentualen Anteil an Übereinstimmung an. Eine Homologie-Angabe erfolgte allerdings nur, wenn eine laut dem Programm „signifikante“ Homologie vorlag.

Bei Sequenzumgebungen ohne eine ausgeprägte Homologie wurde der Anteil an Übereinstimmung durch Abzählen berechnet.

3.7.3. Repetitive Sequenzelemente

Einzelne Elemente in den Bruchpunktumgebungen wurden mit dem Programm der Internetseite <http://genome.ucsc.edu> dargestellt. Die Übereinstimmungen zwischen verschiedenen dieser Elemente und deren genomische Orientierung zueinander wurden mit dem Programm der Internetadresse <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi> untersucht.

Mittels des Programms der Internetadresse <http://repeatmasker.genome.washington.edu> wurden alle beteiligten Introns auf ihren prozentualen Gehalt an den häufigsten repetitiven Elementen untersucht. Dabei handelt es sich um:

- SINEs: Alus, MIRs
- LINEs: LINE1, LINE2, L3/CR1
- LTR-Elemente: MaLRs, ERVL, ERV_KlasseI, ERV_KlasseII
- DNA-Elemente: MER1, MER2
- einfache Repeats

3.7.4. Rekombinations-assoziierte Motive

Mit dem Programm von http://www.bioinformatics.org/sms2/dna_pattern.html wurden jeweils 50bp in 5'- und 3'-Richtung um die Bruchpunktregionen (zusätzlich zu den Bruchpunktregionen) sowie 16x 100bp Vergleichssequenz, die zufällig aus verschiedenen SPAST-Introns ausgewählt worden waren, auf das Auftreten folgender Sequenzmotive untersucht:

- *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats*: GGGGT, TGAGC, GAGCT
- Translin-Bindestelle: GCCCAGCT.

Außerdem wurde mit diesem Programm der AT-Gehalt bestimmt.

3.7.5. Palindrome

Ebenfalls wurden die bereits definierten Bruchpunktumgebungen sowie die Vergleichssequenzen hinsichtlich des Auftretens von Palindromen mit <http://bioweb.pasteur.fr/seqanal/interfaces/palindrome.html> analysiert. Eine maximale Lücke von 100bp zwischen den invertierten Repeats wurde, wie vom Programm vorgeschlagen, zugelassen und als minimale Länge wurden 6bp gewählt.

4. Ergebnisse

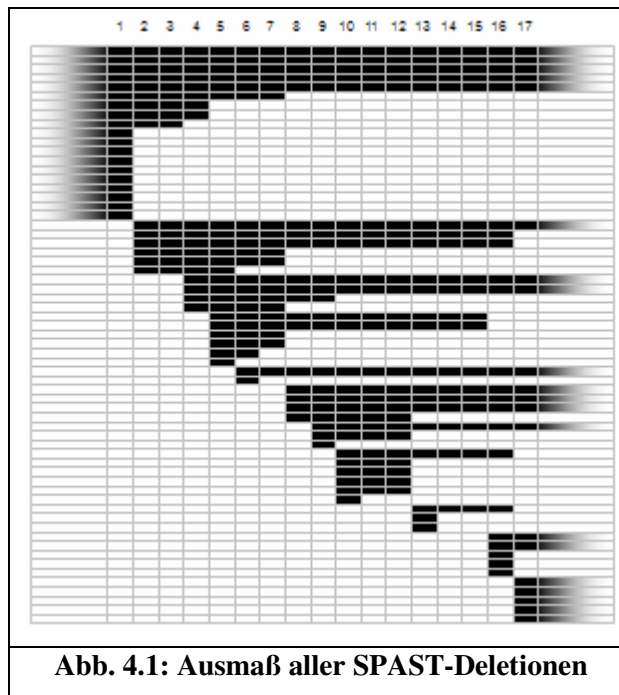
4.1. Analyse der bekannten SPAST-Deletionen

4.1.1. Deletions-Ausmaß

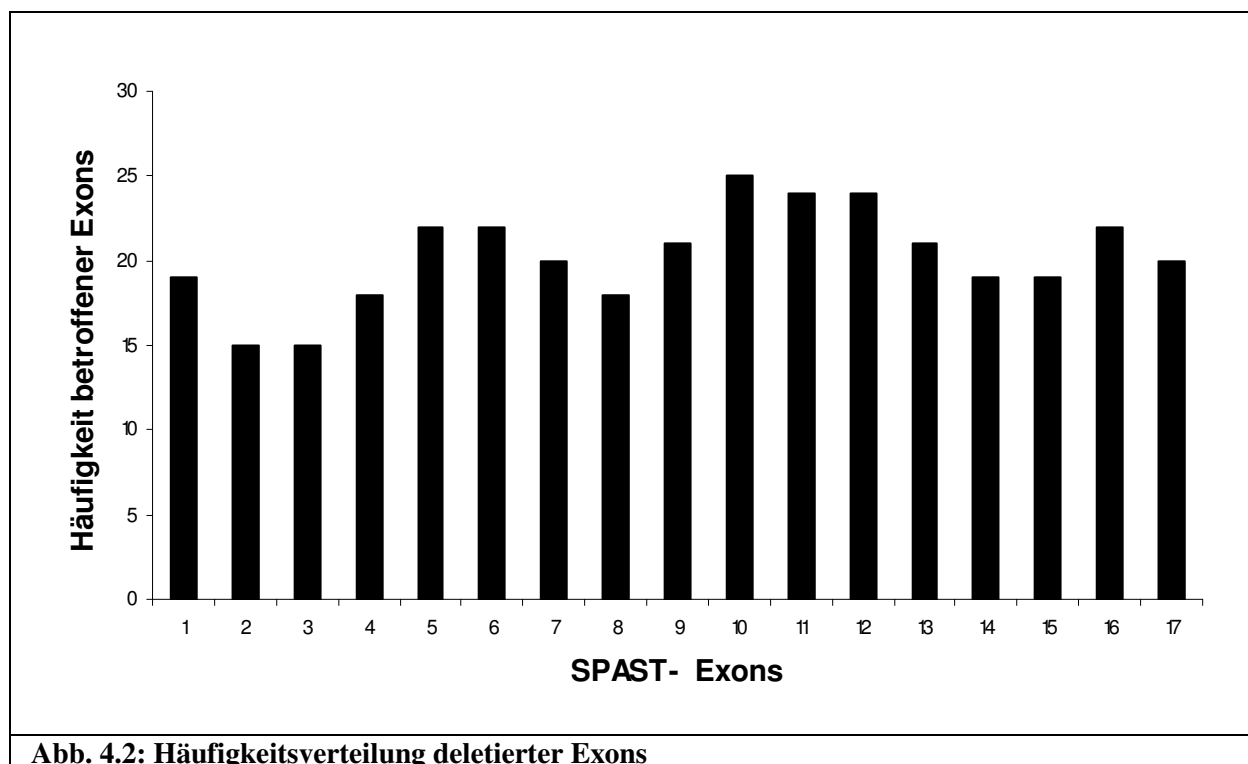
Bei den 63 beschriebenen großen SPAST-Deletionen kommen sowohl Einzel-Exon-Deletionen, Deletionen von Exon-Kombinationen als auch Deletionen, welche die gesamte kodierende Sequenz umfassen, vor. Einzel-Exon-Deletionen haben mit 24/63 Deletionen einen Anteil von 38,1%. Dabei ist das Exon1 am häufigsten betroffen (10/24). Weitere fünf SPAST-Deletionen sind auf das Exon17 und drei Deletionen sind auf das Exon16 begrenzt. Auch Einzel-Exon-Deletionen der Exons 5,6,9,10 und 13 kommen vor. Die Mehrheit der 63 Ereignisse (34/63, was 54,0% entspricht) stellen Deletionen von Exon-Kombinationen in variierendem Ausmaß dar. Damit sind bei der Mehrheit >1 und < 17 Exons deletiert. So sind beispielsweise bei DelEx5_6 nur zwei Exons deletiert, wohingegen bei der DelEx2_17 16 Exons von der Deletion erfasst sind. Die Exon-Kombination DelEx10_12 ist die häufigste dieser Gruppe der SPAST-Deletionen, wobei ihr Anteil mit vier von 34 Ereignissen (11,8%) gering ist. Auch die DelEx8_17 konnte dreimal identifiziert werden und von sieben weiteren Exon-Kombinationen gibt es bereits zwei Deletionsereignisse. Damit dominiert aber keine der Exon-Kombinationen. Alle anderen 13 Deletionen von Exon-Kombinationen sind bis jetzt einmalig und machen 38,2% dieser Gruppe aus. Bei fünf der 63 großen SPAST-Deletionen fehlen alle Exons, so dass jeweils die gesamte kodierende Region vom Deletionsereignis betroffen ist.

29 der 63 Deletionen stellen intragenische Deletionen dar, d.h. dass sich die Deletionen nicht über das 5'-bzw. 3'-Ende des Gens ausdehnen und sowohl Exon1 als auch Exon17 vorhanden sind. Diese Ereignisse machen damit einen Anteil von 46,0% aus.

Dagegen erstrecken sich 34 der 63 SPAST-Deletionen zumindest partiell auch auf Sequenzen außerhalb der kodierenden Region. 19 Ereignisse involvieren Sequenzen, die über das 5'-Ende hinausgehen und 20 Deletionen dehnen sich über das 3'-Ende hinweg aus. Fünf der 34 (partiell) externen Deletionen involvieren sowohl Sequenzen, die über das 5'-Ende hinausgehen als auch das 3'-Ende überragende Regionen. Dies führt dazu, dass die gesamte kodierende SPAST-Sequenz fehlt und damit das gesamte Gen deletiert ist (Abb. 4.1).



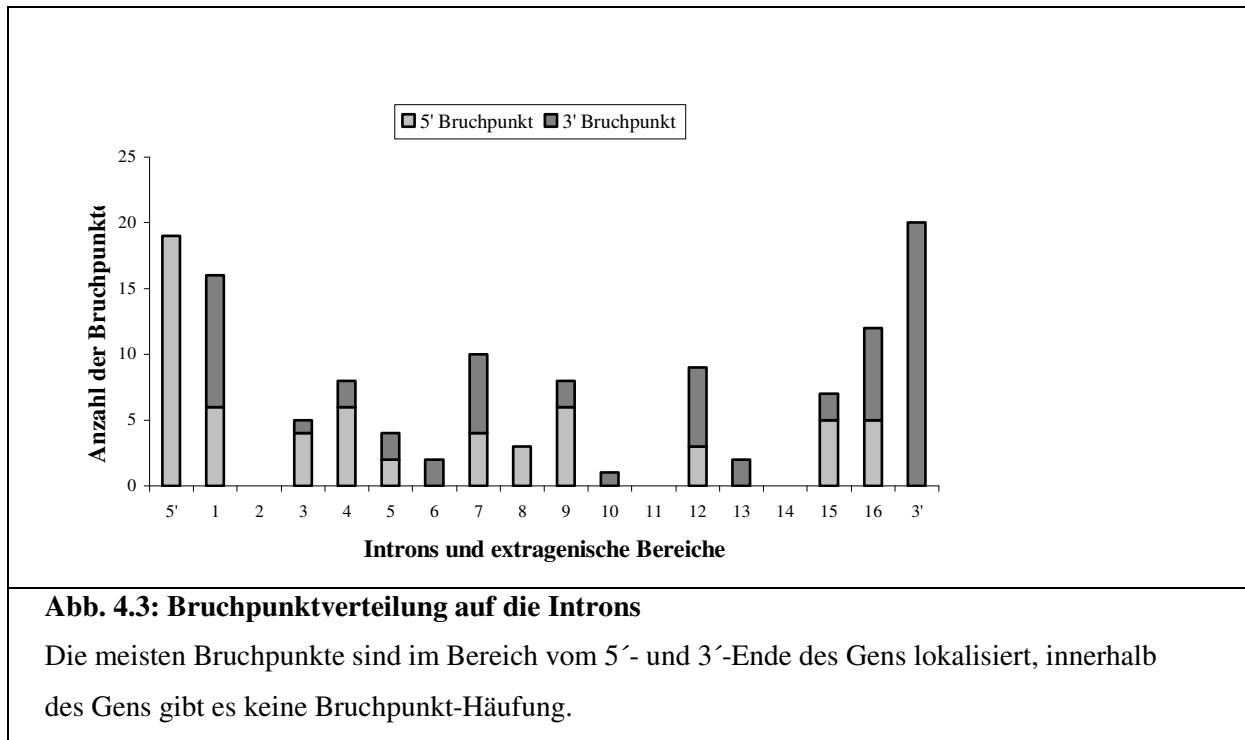
Insgesamt sind alle Exons involviert. Das Exon10 ist am häufigsten betroffen und fehlt bei 25 der 63 großen SPAST-Deletionen (40%). Die Exons mit der geringsten Beteiligung sind sowohl das Exon2 als auch Exon3. Eine Deletion dieser Exons liegt trotzdem in 15 der 63 Fälle vor, was 24% ausmacht. Eine gewisse Häufung lässt sich im Bereich von Exon4-7 und 9-13 vermuten (Abb. 4.2).



4.1.2. Verteilung der Bruchpunkte auf die Introns

Die 63 Deletionsereignisse liefern 126 Bruchpunkte. Dabei liegen 39 der 126 Bruchpunkte (31%) extragenisch, d.h. 5′ vom Exon1 bzw. 3′ vom Exon17. So lassen sich in der 5′-extragenischen Region 19/63 (30%) der 5′-Bruchpunkte finden und in der Region 3′ von Exon17 20/63 (32%) der 3′-Bruchpunkte.

Mehr als 2/3 der Bruchpunkte (87/126, also 69%) sind intragenisch in den SPAST-Introns lokalisiert. Dabei kommen die meisten Bruchpunkte im Intron1 bzw. im Intron16 mit einer Häufigkeit von 16 (sechs 5′- und zehn 3′-Bruchpunkte) bzw. zwölf (fünf 5′- und sieben 3′-Bruchpunkte) vor (Abb.4.3).



Zusammen mit den 5′- und 3′-extragenischen Regionen sind das erste und das letzte Intron fast für die Hälfte aller Bruchpunkt-Lokalisationen verantwortlich. So fallen 25 aller 5′-Bruchpunkte in den 5′-extragenischen Bereich und in das Intron1, was 40% entspricht. 27 aller 3′-Bruchpunkte sind im 3′-extragenischen Bereich sowie im Intron16 lokalisiert, was 43% ausmacht. Weiterhin liegen zehn Bruchpunkte (vier 5′- und sechs 3′-Bruchpunkte) im Intron7, neun Bruchpunkte (drei 5′ und sechs 3′) im Intron12 und jeweils acht Bruchpunkte (jeweils sechs 5′- und zwei 3′-Bruchpunkte) in den Introns 4 und 9. Keine der bekannten großen Deletionen zeigt einen ihrer Bruchpunkte in folgenden Introns: 2, 11 und 14. Alle

weiteren Introns beherbergen einen bis sieben Bruchpunkte ohne eine Häufung in bestimmten Regionen. Aus Abb. 4.3 lässt sich weiterhin auch keine Anreicherung der Introns 2-15 mit Bruchpunkten erkennen.

4.1.3. Zusammenhang zwischen Anzahl der Bruchpunkte/Intron und Intron-Größe

Da allein die Verteilung und die Häufung von Bruchpunkten in bestimmten Introns, ohne die Größe der Introns zu berücksichtigen, keine Auskunft über eine Bruchpunkt-Clusterung oder einen Hotspot liefert, soll nun der Zusammenhang zwischen der Bruchpunkt-Häufung und der Größe der Introns untersucht werden. So macht es einen Unterschied, ob man z.B. in einem kleinen Intron mit nur 140bp, wie es Intron11 ist, vier Bruchpunkte finden würde oder in einem sehr großen Intron mit 23245bp, wie das Intron1.

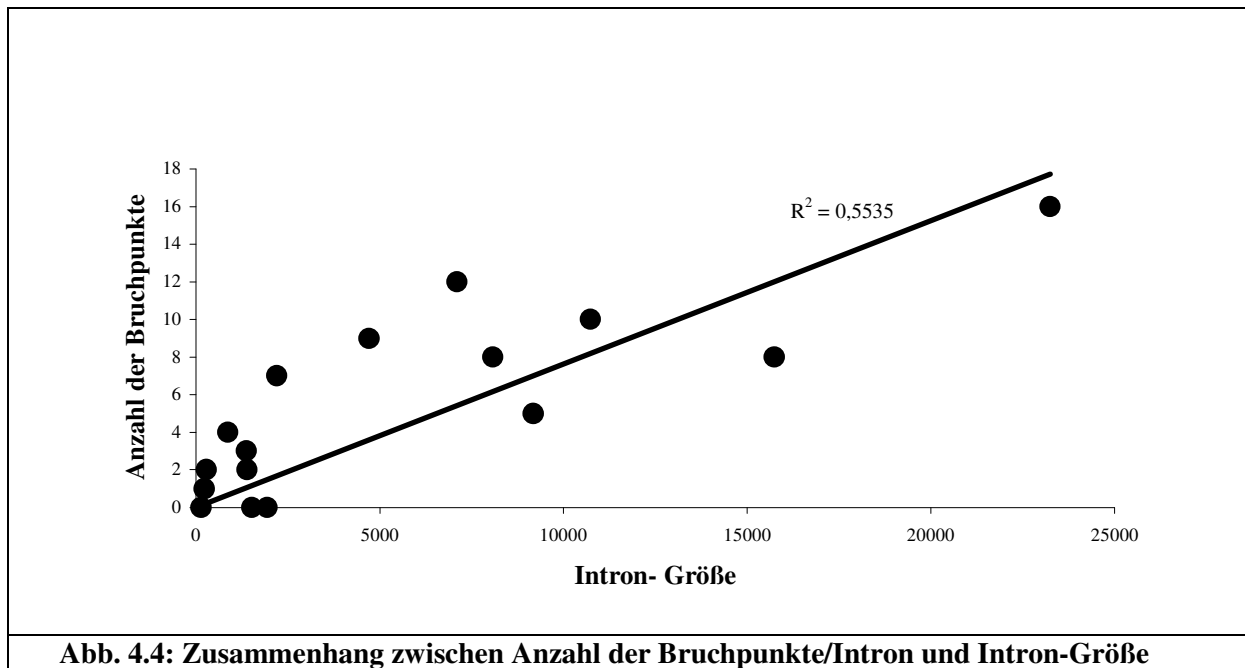
Bei den größten SPAST-Introns, die um 10000bp und größer sind, handelt es sich um das Intron1, Intron4 mit 15746bp, Intron7 mit 10735bp, Intron3 mit 9190bp und Intron9 mit 8083bp. Diese fünf Introns enthalten mehr als 50% aller intragenischen Bruchpunkte und allein das große Intron1 enthält 16/87 der intragenischen Bruchpunkte, was 18% ausmacht. Auch die anderen vier großen Introns beherbergen viele Bruchpunkte. So wurden im Intron4 acht Bruchpunkte gefunden, zehn im Intron7, weitere acht im Intron9 und noch fünf Bruchpunkte im Intron3.

Bei der Betrachtung der Bruchpunkt-reichsten Introns lassen sich die meisten der oben erwähnten großen Introns wieder finden. Dabei enthält das größte Intron (Intron1) auch die meisten Bruchpunkte. Zu dieser Gruppe zählen auch Intron7, 4 und 9. Zwei weitere Introns enthalten noch einen großen Anteil an Bruchpunkten: Intron16 und Intron12 mit zwölf bzw. neun Bruchpunkten, was noch ein Viertel aller intragenischen Bruchpunkte ausmacht. Diese Introns lassen sich mit 7115bp bzw. 4715bp zumindest als Introns mittlerer Größe einstufen. Somit sind sieben von 16 Introns für fast 80% aller intragenischen SPAST-Bruchpunkte verantwortlich.

Die kleinsten SPAST-Introns sind Intron11 mit 140bp Größe und keinem entdeckten Bruchpunkt, Intron10 mit 238bp und einem Bruchpunkt, Intron6, welches 283bp groß ist und zwei bekannte Bruchpunkte beherbergt, und Intron5 mit einer Größe von 876bp enthält immerhin schon vier Bruchpunkte.

In insgesamt drei Introns konnten bisher keine Bruchpunkte gefunden werden: Intron2, 11 und 14. Dabei ist keines dieser Introns größer als 2000bp.

Wie in der Abb. 4.4 zu erkennen ist, lässt sich auch graphisch ein Zusammenhang zwischen der Häufigkeit von Bruchpunkten in einem Intron und der Intron-Größe erkennen. Eine direkte Proportionalität und annähernd eine Linearität kann gezeigt werden. Der Korrelationskoeffizient beträgt 0,5535. Somit verteilen sich viele Bruchpunkte auf große Introns und in kleineren Introns sind nur einige Bruchpunkte zu finden.



4.1.4. Zusammenhang zwischen dem Auftreten von Bruchpunkten und dem Alu-Gehalt

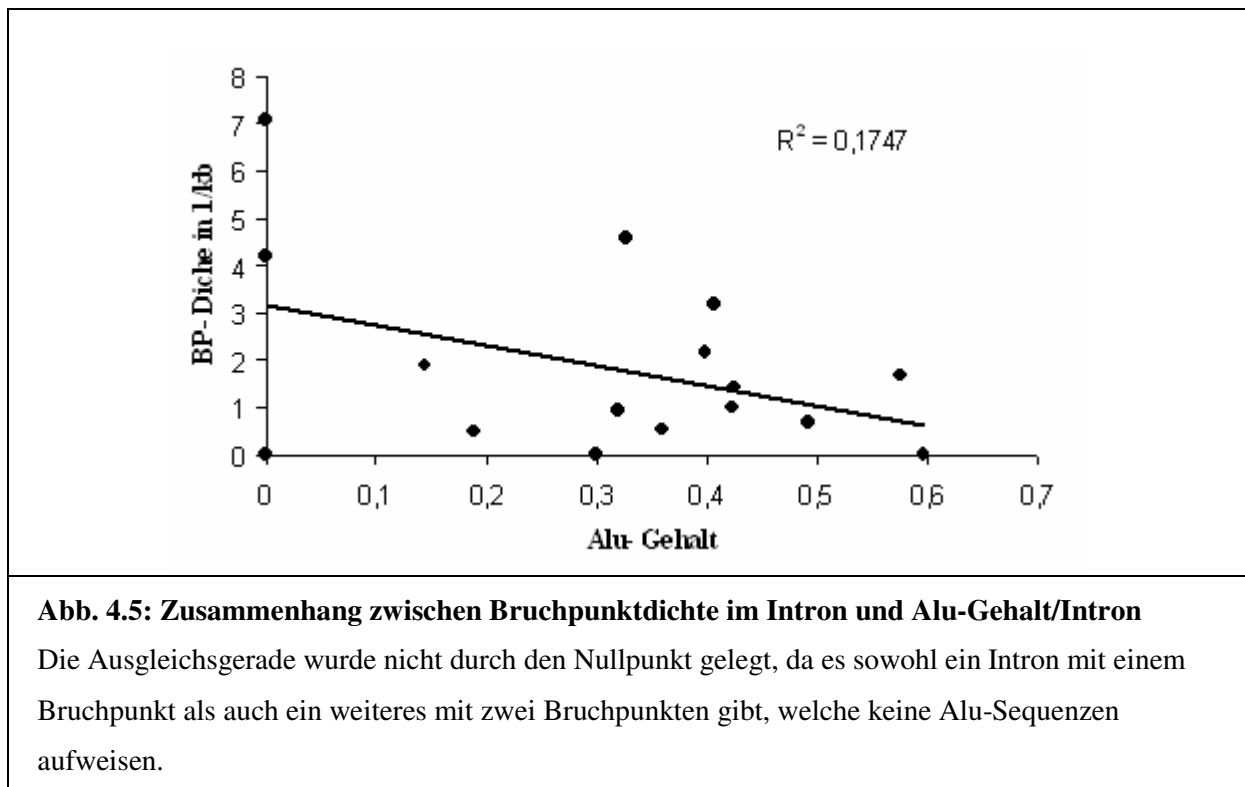
Ein Zusammenhang zwischen dem Alu-Gehalt pro Intron und dem Auftreten von Bruchpunkten soll untersucht werden für die Differenzierung zwischen zufälligen Ereignissen und Sequenz-abhängigen Rekombinationen. Alu-Elemente machen ca. 40% der gesamten SPAST-Sequenz aus. Dennoch gibt es im Alu-Gehalt lokale Unterschiede zwischen den einzelnen Introns. Alu-reiche Introns sind Intron2 mit dem höchsten Alu-Gehalt von 60%, welches aber nicht einen einzigen der bisher entdeckten Deletions-Bruchpunkte beherbergt, das Intron16 (57% Alu-Gehalt) mit zwölf Bruchpunkten ebenso zu den Bruchpunkt-reichen Introns gehörenden DNA-Abschnitten im SPAST-Gen wie das Intron1 mit 16 Bruchpunkten bei einem noch als erhöhten Alu-Gehalt zu bezeichnenden Anteil von 49%.

Anders herum betrachtet, weisen folgende der Bruchpunkt-reichen Introns einen hohen Alu-Gehalt auf: Intron1 und 16 (siehe auch 4.1.3.). Weitere Introns mit vielen Bruchpunkten (Intron7 mit zehn Bruchpunkten und Intron9 mit acht Bruchpunkten) können nur einen mäßigen Alu-Gehalt von 32% bzw. 42% aufweisen. Dagegen zeigen Intron12 und 4, die auch

neun bzw. acht Bruchpunkte beherbergen, nur einen Alu-Gehalt von 14% bzw. 19%, was weit unter dem mittleren Alu-Gehalt im SPAST-Gen liegt.

Die drei Introns, bei denen kein Bruchpunkt bisher identifiziert werden konnte, zeigen sowohl einen hohen und mittleren Alu-Gehalt oder weisen gar kein Alu-Element auf. So weist, wie oben bereits erwähnt, Intron2 zwar den höchsten Alu-Gehalt aller Introns auf, zeigt aber keinen Bruchpunkt. Intron14 verfügt bei einem Alu-Gehalt von 30% über keinen Bruchpunkt. Auch Intron11, ohne Alu-Elemente, weist keinen bekannten Bruchpunkt auf. Allerdings ist das zuletzt genannte Intron auch das kleinste (siehe 4.1.3.). Insgesamt finden sich zwar die meisten Bruchpunkte in Introns mit einem Alu-Gehalt >30%, es gibt aber auch nur wenige, kleine Introns mit einem niedrigeren Alu-Gehalt.

Da in der absoluten Bruchpunkt-Anzahl pro Intron der Größeneffekt der Introns enthalten ist (siehe 4.1.3.), soll nun die Bruchpunktdichte mit dem Alu-Gehalt pro Intron verglichen werden. Introns mit einer hohen Bruchpunktdichte sind die kleinen Introns 4, 6 und 10 mit Bruchpunktdichten von 4,57/kb; 7,07/kb und 4,2/kb. Trotzdem liegen die absoluten Bruchpunkt-Anzahlen nur zwischen einem und vier Bruchpunkten in diesen Introns. Der Alu-Gehalt in diesen Introns macht 0% und bei Intron10 33% aus (siehe Abb. 4.5). Nur Intron15 mit einer noch erhöhten Bruchpunktdichte von 3,17/kb bei sieben absoluten Bruchpunkten weist einen 41%igen Alu-Gehalt auf, der in etwa dem Alu-Gehalt des gesamten SPAST-Gens entspricht.



Auch graphisch lässt sich keine direkte Proportionalität zwischen Alu-Gehalt/Intron und der Bruchpunktdichte finden (Abb.4.5). Stattdessen hat die Ausgleichsgerade sogar einen negativen Anstieg, jedoch mit einem niedrigen Korrelationskoeffizienten. Somit besteht kein Zusammenhang zwischen Alu-Gehalt pro Intron und dem Auftreten von Bruchpunkten.

4.2. Klonierung der zehn Fusionssequenzen

4.2.1. PCR-Produkte

Von allen zehn DNA-Templates konnten genomische Sequenzen amplifiziert werden, die sich durch mehrere Banden für jedes Template in der Gelelektrophorese darstellten. In einzelnen Reaktionen wurde aber auch nur die spezifische Bande amplifiziert.

Beispielhaft zeigt Abb. 4.6 eine Bande von P1, wobei der verwendete r-Primer Int07_r01 nur mit dem f-Primer Int04_f09 ein Ergebnis liefert, während die anderen forward-Primer zu weit vom reverse-Primer Int07_r01 entfernt lokalisiert sind.

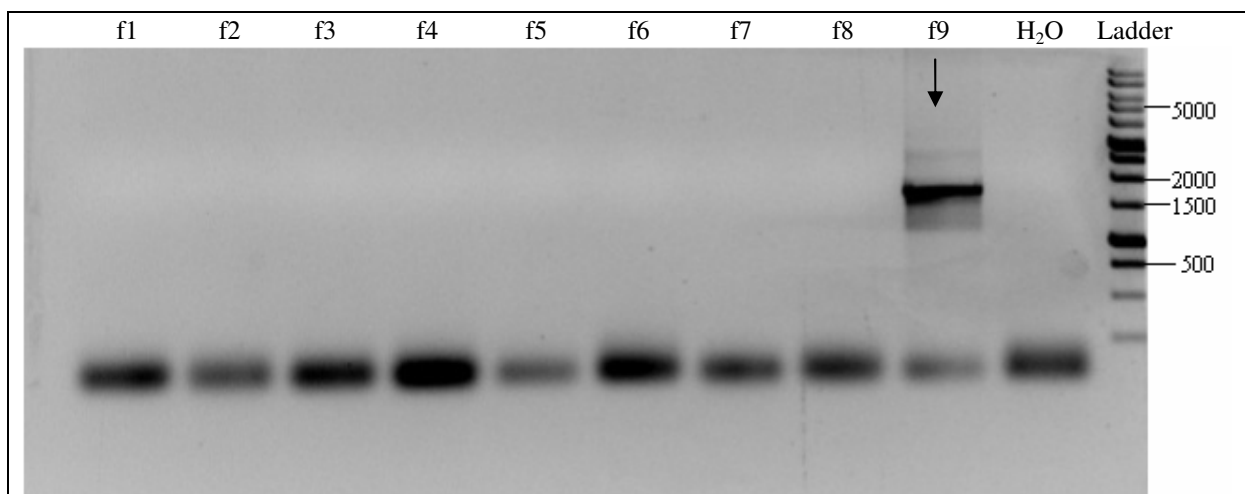


Abb. 4.6: PCR von P1

Nur die Kombination Int04_f09 vs. Int07_r01 liefert bei P1 eine Bande mit einer Größe von ca. 1700bp.

Es gab auch Templates, bei denen sich keine der dargestellten Banden im Nachhinein als spezifisch herausstellte. So ergab die Amplifikation von P4 kein erwünschtes Produkt des SPAST-Gens, sondern nur unspezifische Amplifikationen von anderen Chromosomen.

Des Weiteren wurde bei P9 nur eine Bande erhalten mit den Primern Int15_f02 vs. Intr16_r02 von folgender Größe: ca. 2000-3000bp. Diese Bande entspricht dem Abstand der gewählten

Primer, also auch dem Wildtyp. Dessen Amplifikation war möglich durch eine Deletion von nur Exon16, also auch einer geringen Distanz der Vorwärts- und Rückwärts-Primer schon im gesunden Allel. Eine Deletions-typische Bande konnte bei P9 somit nicht dargestellt werden.

Es wurden außerdem auch Banden erhalten, die von Anfang an als unspezifisch deklariert werden konnten und somit gar nicht erst ausgeschnitten wurden.

Weiterhin gab es Ansätze, bei denen mehrere Primer-Kombinationen eine Bande in der Gelelektrophorese lieferten, ohne dass diese als unspezifisch gewertet werden konnten. So ergab die Primer-Kombination Int07_r05 vs. Int09_f09 bei der Amplifikation von P7 eine ca. 0,8kb große Bande und auch die Kombination aus dem benachbarten Rückwärts-Primer Int07_r06 mit demselben Vorwärts-Primer dann eine ca. 2kb große Bande (siehe Abb. 4.7).

Dabei betrug der Abstand zwischen den ausgewählten reverse-Primern 1,2kb.

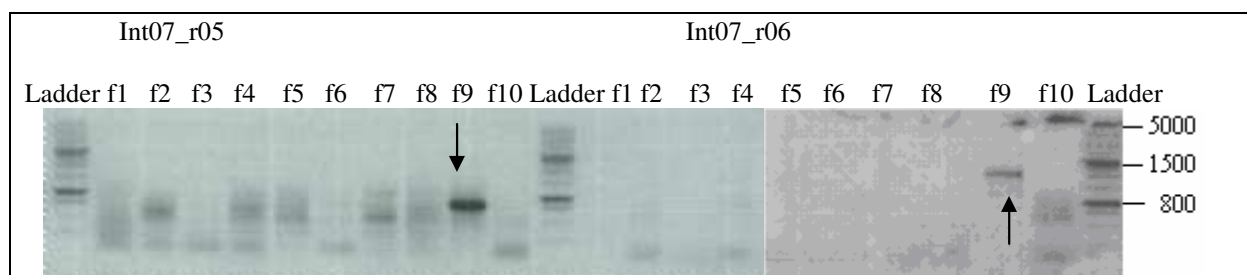


Abb. 4.7: PCR von P7

Beim Template P7 ergab Int09_f09 vs. Int07_r05 eine Bande sowie Int04_f09 vs. Int07_r06.

Bei einzelnen Templates, die nur die Deletion eines Exons aufweisen, konnten im selben Ansatz sowohl das gesunde Allel (Wildtyp) als auch das mutierte Allel amplifiziert werden (siehe Abb. 4.8).

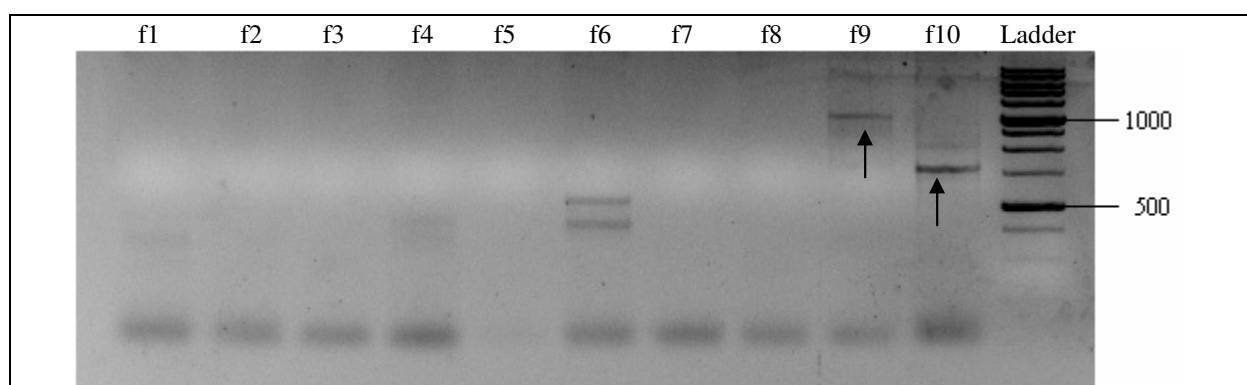


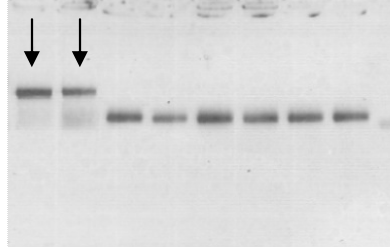
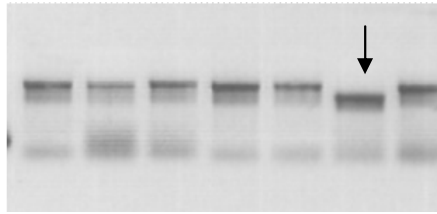
Abb. 4.8: PCR von P3

Int04_f09 vs. Int07_r01 ergab die Wildtyp-Bande und Int04_f10 vs. Int07_r01 das Allel mit der Deletion.

Dabei ergab die Primer-Kombination Int04_f09 vs. Int07_r01 die Amplifikation des Normaltyps mit einer Größe von ca. 3000bp und die Kombination aus Int04_f10 vs. Int07_r01 das gesuchte deletierte Allel von P3. Die Größe der Bande ergab ca. 1000bp bei einer Deletions-Größe von ca. 800bp und einem Abstand zwischen Int04_f09 und Int04_f10 von 900bp.

4.2.2. Kolonie-PCR-Produkte

In einigen Kolonie-PCR-Ansätzen konnten keine Banden erhalten werden. Weiterhin wurden in einigen Fällen Banden sichtbar, die von der Größe nicht zu den initialen PCR-Produkten passten und dann nicht weiter bearbeitet wurden. Insgesamt wurden aber von allen initialen PCR-Produkten auch Banden mit passender Größe durch die Kolonie-PCR dargestellt. Dies entspricht dem Nachweis, dass die Klonierung erfolgreich war und der gewünschte DNA-Abschnitt im amplifizierten Abschnitt enthalten ist, was das Ziel dieses Verfahrens darstellt. Als Beispiel zeigt eine Kolonie-PCR vom Template P7 acht Banden zweier unterschiedlicher Größen (siehe Abb. 4.9a). Dabei entsprechen nur die ersten beiden Banden der erwarteten Größe von 0,8kb, während die anderen Banden kleiner sind und damit als nicht passend einzustufen sind.

| | |
|--|---|
|  |  |
| <p>Abb. 4.9a: Kolonie-PCR von P7: Die ersten beiden Banden entsprechen einer Größe von 0,8kb.</p> | <p>Abb. 4.9b: Kolonie-PCR von P9: Nur die sechste Bande entspricht der erwarteten Größe.</p> |

Auch bei der Kolonie-PCR von P9 entsprechen sechs von sieben Banden nicht der erwarteten Größe, verglichen mit der initialen PCR. Nur die vorletzte Bande kann damit als spezifisches initiales PCR-Produkt gezählt werden, da es sich mit der erwarteten Größe von 0,8kb darstellte und somit dem gesuchten DNA-Anschnitt entspricht (Abb. 4.9b).

4.2.3. Sequenzen

Von allen zehn DNA-Templates wurden klonierte PCR-Produkte sequenziert. Dabei gab es viele Sequenzen, die nicht der Fusionssequenz entsprachen, sondern Amplifikationen anderer genomischer Bereiche außerhalb des Chromosoms 2 darstellten. Außerdem wurde auch in einigen Fällen der Wildtyp sequenziert, was daran zu erkennen war, dass beim Ablesen und gleichzeitigem Vergleich mit der „Normalsequenz“ eine komplette Übereinstimmung bestand.



Abb. 4.10:
Sequenz von P10

Von acht der zehn initialen Templates konnten dann die spezifischen Fusionsregionen der Deletionen sequenziert werden. Somit stellte sich erst an dieser Stelle heraus, dass nur acht der zehn Templates spezifisch PCR amplifiziert wurden.

Als Beispiel-Sequenz ergibt sich, von „oben“ gelesen, folgende Sequenz von P10: ttttttttttttttttaatcgaccgcgaaccaccgcacgtggacatcagggtc gatgagccctccgactccgctc.

Bis dahin gibt es in umgekehrter Reihenfolge (siehe Pfeil in Abb. 4.10) eine Übereinstimmung mit der Sequenz aus dem Intron16. Hier geht es weiter mit: tcagaacgagacaacggg und im Intron16 folgt die Sequenz: ctcttagtgaactgggtca, also fehlen ab hier (Pfeil) Nukleotide vom Intron16. Wenn man dann von „unten“ abliest, so ergibt sich folgende Sequenz: agtaagccgagcactgcagcctgggcaacagagcaagactctg.

Bis zu dieser Stelle gibt es eine Übereinstimmung mit der Sequenz aus dem Intron15. Weiter geht es in der Fusionssequenz hier mit cctcagcctcccgagta, während im Intron15 tctcaaaaaaaaaaaaaa folgt und ab hier (Pfeil) Nukleotide vom Intron15 fehlen.

Somit ergibt sich folgende Fusionssequenz:

agtaagccgagcactgcagcctgggcaacagagcaagactctgcctcagcctcccgagtagctggg actacaggtgcacgcca, wobei der erste Sequenzabschnitt der „unten“

abgelesenen und der genomischen Sequenz aus dem Intron15 entspricht und der zweite Abschnitt (ab blau markierter Sequenz) der umgekehrt von „oben“ abgelesenen Sequenz und der aus dem Intron16 entspricht. Die blau-schattierten Nukleotide entsprechen der Überlappung, also der Sequenz, die sowohl im Intron15 als auch im Intron16 so vorkommen. Sie entsprechen der Bruchpunktregion, wobei der Bruchpunkt nicht auf ein Nukleotid genau bestimmt werden kann. Somit ergeben sich die acht Fusionssequenzen aller Templates in der Abb. 4.11.

[illegible]

Abb. 4.11: Alle Fusionssequenzen

Obere Reihe: Intronische Sequenz um den 5'-Bruchpunkt

| Mittlere Reihe: | Fusionssequenz |
|-----------------|----------------|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 7 |
| 8 | 8 |
| 9 | 9 |
| 10 | 10 |
| 11 | 11 |
| 12 | 12 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 18 | 18 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 24 | 24 |
| 25 | 25 |
| 26 | 26 |
| 27 | 27 |
| 28 | 28 |
| 29 | 29 |
| 30 | 30 |
| 31 | 31 |
| 32 | 32 |
| 33 | 33 |
| 34 | 34 |
| 35 | 35 |
| 36 | 36 |
| 37 | 37 |
| 38 | 38 |
| 39 | 39 |
| 40 | 40 |
| 41 | 41 |
| 42 | 42 |
| 43 | 43 |
| 44 | 44 |
| 45 | 45 |
| 46 | 46 |
| 47 | 47 |
| 48 | 48 |
| 49 | 49 |
| 50 | 50 |
| 51 | 51 |
| 52 | 52 |
| 53 | 53 |
| 54 | 54 |
| 55 | 55 |
| 56 | 56 |
| 57 | 57 |
| 58 | 58 |
| 59 | 59 |
| 60 | 60 |
| 61 | 61 |
| 62 | 62 |
| 63 | 63 |
| 64 | 64 |
| 65 | 65 |
| 66 | 66 |
| 67 | 67 |
| 68 | 68 |
| 69 | 69 |
| 70 | 70 |
| 71 | 71 |
| 72 | 72 |
| 73 | 73 |
| 74 | 74 |
| 75 | 75 |
| 76 | 76 |
| 77 | 77 |
| 78 | 78 |
| 79 | 79 |
| 80 | 80 |
| 81 | 81 |
| 82 | 82 |
| 83 | 83 |
| 84 | 84 |
| 85 | 85 |
| 86 | 86 |
| 87 | 87 |
| 88 | 88 |
| 89 | 89 |
| 90 | 90 |
| 91 | 91 |
| 92 | 92 |
| 93 | 93 |
| 94 | 94 |
| 95 | 95 |
| 96 | 96 |
| 97 | 97 |
| 98 | 98 |
| 99 | 99 |
| 100 | 100 |

Untere Reihe: Intronische Sequenz um den 3'-Bruchpunkt

! Übereinstimmung der 5'- und 3'-Bruchpunktregionen

Absolute Homologie

| schwarz | Vorhandene Sequenz |
|---------|--------------------|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 7 |
| 8 | 8 |
| 9 | 9 |
| 10 | 10 |
| 11 | 11 |
| 12 | 12 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 18 | 18 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 24 | 24 |
| 25 | 25 |
| 26 | 26 |
| 27 | 27 |
| 28 | 28 |
| 29 | 29 |
| 30 | 30 |
| 31 | 31 |
| 32 | 32 |
| 33 | 33 |
| 34 | 34 |
| 35 | 35 |
| 36 | 36 |
| 37 | 37 |
| 38 | 38 |
| 39 | 39 |
| 40 | 40 |
| 41 | 41 |
| 42 | 42 |
| 43 | 43 |
| 44 | 44 |
| 45 | 45 |
| 46 | 46 |
| 47 | 47 |
| 48 | 48 |
| 49 | 49 |
| 50 | 50 |
| 51 | 51 |
| 52 | 52 |
| 53 | 53 |
| 54 | 54 |
| 55 | 55 |
| 56 | 56 |
| 57 | 57 |
| 58 | 58 |
| 59 | 59 |
| 60 | 60 |
| 61 | 61 |
| 62 | 62 |
| 63 | 63 |
| 64 | 64 |
| 65 | 65 |
| 66 | 66 |
| 67 | 67 |
| 68 | 68 |
| 69 | 69 |
| 70 | 70 |
| 71 | 71 |
| 72 | 72 |
| 73 | 73 |
| 74 | 74 |
| 75 | 75 |
| 76 | 76 |
| 77 | 77 |
| 78 | 78 |
| 79 | 79 |
| 80 | 80 |
| 81 | 81 |
| 82 | 82 |
| 83 | 83 |
| 84 | 84 |
| 85 | 85 |
| 86 | 86 |
| 87 | 87 |
| 88 | 88 |
| 89 | 89 |
| 90 | 90 |
| 91 | 91 |
| 92 | 92 |
| 93 | 93 |
| 94 | 94 |
| 95 | 95 |
| 96 | 96 |
| 97 | 97 |
| 98 | 98 |
| 99 | 99 |
| 100 | 100 |

grau **Deletierte Sequenz**

4.3. In silico-Analyse der Bruchpunktregionen

4.3.1. Homologie an den Fusionspunkten

Die Sequenzierung der Fusionssequenzen ergab eine Überlappung der Sequenzen von 5' und 3'. Durch diese Homologie an den jeweiligen Verbindungen kann der definitive Bruchpunkt nicht vorausgesagt werden (siehe Abb. 4.11). Mindestens 3bp und maximal 14bp stimmen überein.

Dabei befindet sich bei drei der acht Fusionssequenzen eine 3bp-Homologie. Es handelt sich hierbei um die Nukleotidfolge aaa bei P3, ttt bei P8 und ctg bei P10. Somit besteht bei zwei von drei der 3bp-langen Homologien ein Nukleotid-Repeat. Zwei von acht Sequenzen zeigen eine 4bp-lange exakte Übereinstimmung. Dies ist bei P1 atga und bei P6 ttcc. An einem Fusionspunkt sind 6bp homolog und zwar tgagcc bei P5. P7 zeigt eine 7bp-lange Homologie am Fusionspunkt: ttctatc. Bei P2 findet sich eine Überlappung von 14bp, bei der es sich um die Sequenz agtagctgggatta handelt.

Insgesamt überwiegen kleine Homologien, bei denen es sich um eine Mikrohomologie ≤ 7 bp handelt, mit sieben von acht gegenüber einer nur einmal gefundenen großen Übereinstimmung.

4.3.2. Homologie in den Bruchpunktumgebungen

Da es sich in folgender Betrachtung um die 50bp-Umgebung jeweils in 5'- und 3'-Richtung von den Bruchpunktregionen aus handelt, werden diese zur Vereinfachung als „Bruchpunkte“ bezeichnet, auch wenn diese nicht exakt definiert werden können.

Beim Vergleich der Sequenzumgebungen von 5'- und 3'-Bruchpunkt ergibt sich die niedrigste Homologie mit 21% bei P10. Die höchste Homologie liegt bei P2 mit 83%. Auch die Fusionssequenz von P5 zeigt eine hohe Übereinstimmung von insgesamt 63%, die im Bereich von Basenpaar 6-70 sogar einen Bereich mit 84%iger Homologie aufweist und damit der Größenordnung der Homologie von P2 entspricht. Dabei schließt der Abschnitt mit großer Übereinstimmung auch den Fusionspunkt ein. Somit zeigen zwei von acht der Fusionssequenzen eine ausgeprägte Homologie.

Dagegen können sechs von acht der analysierten Sequenzen nur eine geringe Übereinstimmung aufzeigen, die zwischen 21-34% und im Mittel bei 28,5% liegt. Es handelt

sich dabei im Einzelnen um 21% bei P10, 32% bei P3, 29% bei P1, 32% bei P6, 34% bei P7 und 23% bei P8. Diese Übereinstimmung liegt damit um 25%, was gerade dem Zufall entspricht, also der Wahrscheinlichkeit eines Nukleotids, den homologen „Partner“ beim Vergleich der beiden Bruchpunktumgebungen zu finden und wurde vom Programm der Internetadresse <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi> auch als „nicht signifikante Homologie“ eingestuft.

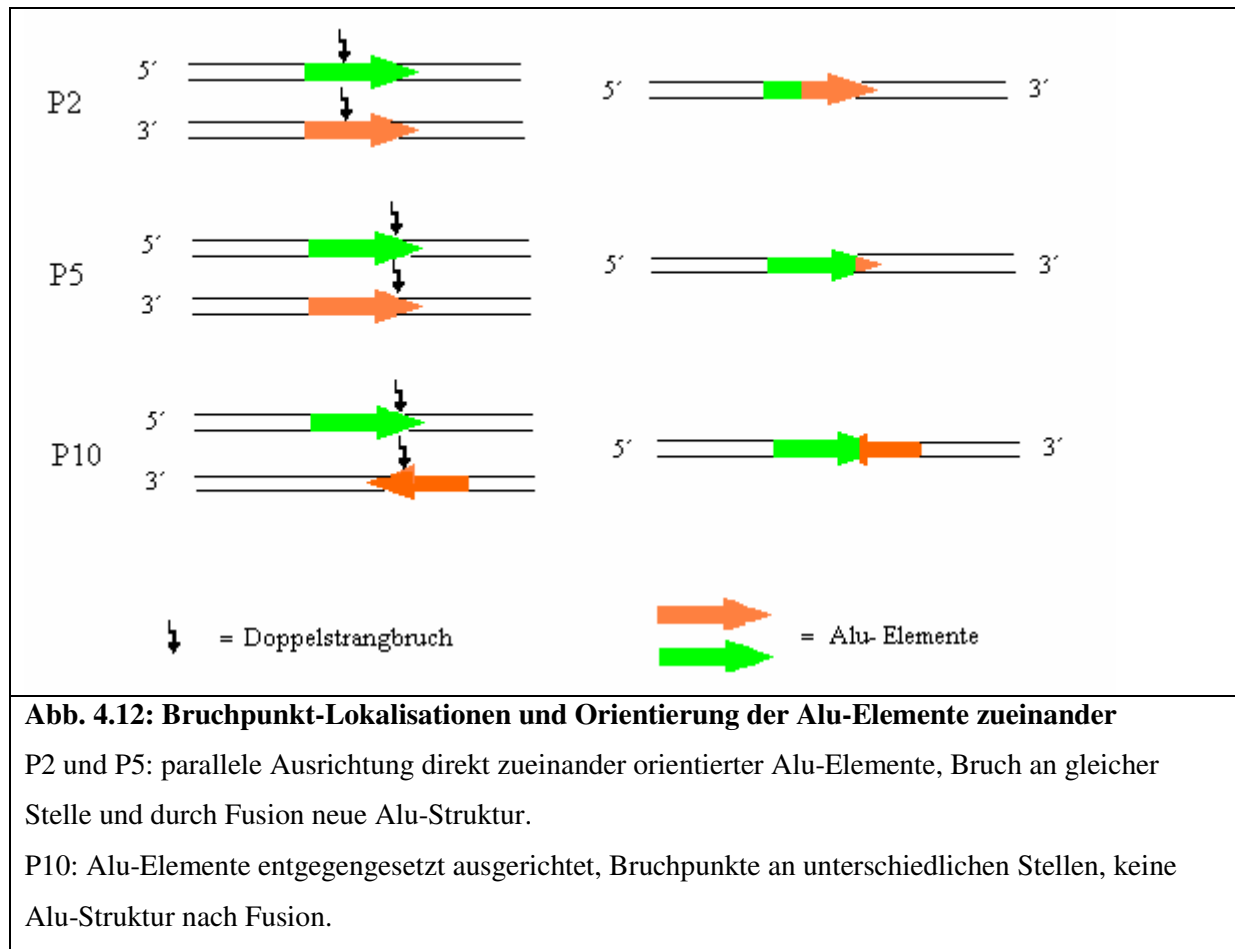
4.3.3. Strukturelle Besonderheiten in den Bruchpunktumgebungen

4.3.3.1. Repetitive Sequenzelemente

Insgesamt befinden sich sieben der 16 Deletions-Bruchpunkte innerhalb eines Alu-Elements. Dies entspricht 44% und liegt damit in der Größenordnung des Alu-Gehalts im gesamten SPAST-Gen (ca. 40%). Bei drei von acht Ereignissen sind gleich beide Bruchpunkte in einem solchen repetitiven Element lokalisiert.

So liegt bei P2 der 5'-Bruchpunkt ca. zwischen dem ersten und zweiten Drittel des AluSx und der 3'-Bruchpunkt an korrespondierender Stelle in einem AluSq-Element. AluSx und AluSq stimmen dabei in 78-85% überein und liegen in direkter Orientierung zueinander vor. Auch P5 hat beide Bruchpunkte an übereinstimmender Lokalisation innerhalb der beteiligten Alu-Elemente AluSx um den 5'-Bruchpunkt und AluY um den 3'-Bruchpunkt. Dabei sind beide Bruchpunkte ca. 15bp 5' vom 3'-Ende der Alu-Elemente zu finden. Es handelt sich hierbei ebenfalls um direkt zueinander orientierte Alu-Elemente mit 80%iger Übereinstimmung. Eine weitere Fusionssequenz (P10) weist beide Bruchpunkte in Alu-Elementen auf. Dabei liegt der 5'-Bruchpunkt ca. 30bp vom 3'-Ende des AluSg entfernt und der 3'-Bruchpunkt dagegen zwischen dem ersten und zweiten Drittel des AluSc, also an nicht korrespondierender Lokalisation. Die beiden beteiligten Alu-Elemente zeigen zwar 84-86%ige Homologiebereiche, allerdings nur in umgekehrter Orientierung zueinander (Abb. 4.12).

Nur einen Bruchpunkt innerhalb eines Alus zeigt P3. Der 5'-Bruchpunkt liegt in einem AluSp-Element, während der 3'-Bruchpunkt sich in nicht repetitiver Sequenz befindet.



Neun der 16 Deletions-Bruchpunkte liegen in nicht repetitiven Sequenzen, wobei sich aber bei drei dieser neun Bruchpunkte in unmittelbarer Nähe (maximal 40bp Entfernung) weitere Alu-Elemente befinden. Dies sind AluSc und AluSx 5' und 3' des 5'-Bruchpunktes von P1, AluSx 5' vom 5'-Bruchpunkt von P6 und AluJb 5' des 5'-Bruchpunktes von P7, wobei jeweils der 3'-Bruchpunkt in nicht repetitiver Sequenz zu finden ist. Am häufigsten befindet sich ein AluSx-Element in den Bruchpunktumgebungen, wobei jeweils zwei Bruchpunkte innerhalb und zwei Bruchpunkte in der Umgebung eines solchen Elements lokalisiert sind (Tabelle 4.1).

Zusätzlich zu der räumlichen Assoziation zu Alu-Elementen weisen zwei von 16 Bruchpunkten in unmittelbarer Umgebung LINE-Strukturen auf. Diese zwei Bruchpunkte stammen vom selben Ereignis, der 5'- und 3'-Bruchpunkt von P8. So zeigt der 5'-Bruchpunkt in seiner Umgebung 5' ein HAL1B und 3' ein L1PA16, während sich 3' des 3'-Bruchpunktes ein L1MC befindet. Diese Elemente lassen aber beim Vergleich keine ausgeprägte Homologie feststellen und gehören nicht zur selben Subfamilie. Die beiden Bruchpunkte dieses Ereignisses liegen in Introns mit einem ausgeprägten LINE-Reichtum. So weist das Intron12

einen LINE-Gehalt von 61% und das Intron13 einen von 47% auf. Andere SPAST-Introns haben dagegen einen geringeren LINE-Gehalt: neun von 16 Introns weisen gar keine LINE-Strukturen auf und mit Ausnahme von Intron4, welches einen LINE-Gehalt von 48% hat, zeigen selbst LINE-haltige Introns nur einen 7-35%igen LINE-Gehalt.

| Familie | Deletion | 5'Sequenzelement | Bezug BP zum Element | 3'Sequenzelement | Bezug BP zum Element |
|---------------------|----------|--------------------------|--|-----------------------|---|
| FSP530 (P2) | Ex5_7 | AluSx | Bruch nach ca. 1/3 (100bp) | AluSq | BP nach ca. 1/3 (100 bp) |
| FSP157 (P5) | Ex9 | AluSx | BP ca. 15bp vom 3'-Ende des AluSx entfernt | AluY | BP ca. 15bp vom 3'-Ende des AluY entfernt |
| | | | | | |
| FSP803 (P10) | Ex16 | AluSg | BP ca. 30bp vom 3'-Ende des AluSg entfernt | AluSc | Bruch nach ca. 1/3 |
| FSP167 (P3) | Ex6 | AluSp | BP im Bereich des 3'-Endes von AluSp | Unique | |
| | | | | | |
| FSP742 (P1) | Ex5_6 | AluSc/AluSx in Umgebung | BP 25-28bp 3' vom AluSc und 40-43bp 5' vom AluSx | Unique | |
| FSP235 (P6) | Ex9_12 | AluSx in Umgebung | BP 27-31bp 3' vom AluSx | Unique | |
| FSP335 (P7) | Ex10_16 | AluJb in Umgebung | BP 29-36bp 3' vom AluJb | Unique | |
| FSP333 (P8) | Ex13 | HAL1B/L1PA16 in Umgebung | BP 6-9bp 3' vom HAL1B und 27-30bp 5' vom L1PA16 | L1MC (L1) in Umgebung | BP 17-20bp 5' vom L1MC |

Tab. 4.1: Sequenzelemente in den Bruchpunktregionen

4.3.3.2. Rekombinations-assoziierte Motive

In den 16 Bruchpunktumgebungen (50bp 5' und 50bp 3') lassen sich sechs bekanntermaßen mit Mutationen und Rekombinationen assoziierte Sequenzmotive, bei denen es sich um *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats* handelt, finden. Dies sind bei P2 ein GGGGT-Motiv, bei P3 TGAGC, bei P5 TGAGC, zwei TGAGC-Motive bei P6 und ein TGAGC bei P8. Lokalisiert sind diese Motive in einer Entfernung von 20-50bp zu den Bruchpunkten. Dagegen weisen die Sequenzen von P1, P7 und P10 kein solches Motiv auf. Am häufigsten wurde damit TGAGC identifiziert und es macht fünf der sechs gefundenen Motive aus. Bei der Analyse von 16 verschiedenen 100bp-Sequenzen, die zufällig aus den

SPAST-Introns ausgewählt wurden, konnten 13 dieser Motive entdeckt werden. Dabei handelt es sich ebenfalls in der Mehrzahl der Fälle (7/13) um das Motiv TGAGC. Außerdem wurden vier GAGCT-Motive und zwei GGGGT-Motive gefunden. Somit sind ca. doppelt so viel *Immunglobulin-Heavy-Chain-Class-Switch-Repeats* in den zufälligen Sequenzen präsent wie in den Bruchpunktregionen. Allerdings weist die Hälfte der zufällig ausgesuchten Sequenzen auch kein solches Motiv auf. Im gesamten SPAST-Gen konnten 545 solcher Motive identifiziert werden. Das bedeutet bei einer Gesamtgröße des Gens von ca. 90000bp, dass im Durchschnitt ein Motiv alle 167bp vorhanden ist. Man würde dann in 1600bp (16 Bruchpunkte mit 50bp jeweils 5' und 3') Sequenz zehn *Immunglobulin-Heavy-Chain-Class-Switch-Repeats* erwarten. Da hier aber nur sechs solcher Motive gefunden wurden, kann man nicht von einer Häufung der Motive in den Bruchpunktumgebungen ausgehen.

Ein zweites untersuchtes Motiv in den Bruchpunktumgebungen sind Translin-Bindestellen-Motive. Dabei wurde das Motiv GCCCAGCT bei P2 in einer Entfernung von 17bp vom Fusionsbereich gefunden. Alle weiteren Fusionssequenzen weisen kein solches Motiv auf.

4.3.3.3. Palindrome

Es konnten in allen acht Fusionssequenzen aus je 2x 100bp-Bruchpunktumgebung (sowohl 5'- als auch 3'-Umgebung) invertierte Repeats (Palindrome) mit einer maximal zugelassenen Lücke von 100bp zwischen den Repeats gefunden werden. Dabei handelt es sich um insgesamt 22 solcher Sequenzen, wovon 18 aus jeweils 6bp, drei aus 7bp und ein Palindrom aus 8bp Sequenz besteht.

Während P6, P5 und P2 (8bp-Repeat) nur eine palindromische Sequenz aufweisen können, verfügen alle anderen Fusionen über mindestens zwei Palindrome. Dabei besitzen 5'- und 3'-Bruchpunktumgebung von P3 die meisten invertierten Repeats mit fünf 6bp-Repeats und einem 7bp-Repeat. Außer bei P5 ist bei allen ein Repeat direkt in der Bruchpunktregion lokalisiert und weist damit eine Überlappung sowohl mit der Mikrohomologie von P1, P3, P6, P7, P8 und P10 als auch mit der 14bp-Homologie im Bereich der Fusion von P2 auf.

In den Vergleichssequenzen von 16x 100bp konnten allerdings auch 20 palindromische Sequenzen identifiziert werden. Dabei handelt es sich um 18 6bp-Repeats, ein 8bp- und ein 9bp-Repeat. Dies entspricht einer fast identischen Anzahl an gefundenen Palindromen beim Vergleich der Bruchpunktumgebungen mit den zufällig gewählten Sequenzen.

4.3.3.4. AT-Reichtum

Die Untersuchung des präferenziellen Auftretens bestimmter Nukleotide in den 100bp-Bruchpunktumgebungen zeigt ein Überwiegen von A und T in den acht Fusionssequenzen. Im Mittel weisen die acht Fusionssequenzen einen AT-Gehalt von 61% auf. Dabei liegt der höchste Gehalt bei P7 mit 69% und der niedrigste bei P2 mit 49%, was bei P2 noch der statistischen Wahrscheinlichkeit entsprechen würde. Alle weiteren Bruchpunktumgebungen zeigen eine AT-Anreicherung von über 50%. Einzelne Bruchpunktumgebungen weisen sogar einen AT-Reichtum von über 75% auf. So liegt der 3'-Bruchpunkt von P1 in einer von 78%iger AT-Sequenz umgebenen Region.

Ebenfalls eine AT-Anreicherung in Form von 17-29bp langen A- oder T-Strecken zeigen die umgekehrt zueinander orientierten Alu-Elemente, die die Bruchpunkte von P10 enthalten.

Außerdem befinden sich bei zwei von drei der 3bp-Mikrohomologien dreifache Wiederholungen eines Nukleotids. Dabei handelt es sich um aaa als Mikrohomologie bei P3 und ttt als Übereinstimmung bei P8. Einfache C- oder G-Repeats direkt am Fusionspunkt kommen bei den hier analysierten Deletionsereignissen nicht vor.

Wenn man aber die bereits erwähnten, zufällig ausgesuchten Vergleichssequenzen auf ihren AT-Gehalt prüft, so zeigen die 16 analysierten Sequenzen, ähnlich dem der Fusionssequenzen, einen Mittelwert von 60%igem AT-Gehalt. Bis auf zwei der 16 zufälligen Sequenzen, deren AT-Gehalt bei nur 38 bzw. 41% liegt, weisen alle anderen Sequenzen Werte >50% auf und der höchste ermittelte Gehalt liegt sogar bei 80%. Insgesamt liegt der AT-Gehalt im SPAST-Gen bei 59,8%, was dem Mittelwert der insgesamt 1600bp Bruchpunktumgebung mit 61% sehr nah kommt. Also kann nicht von einem erhöhten AT-Gehalt in den Bruchpunktumgebungen ausgegangen werden und auch nicht vom AT-Gehalt in einer bestimmten Region auf eine Bruchpunkt-Häufung geschlossen werden.

5. Diskussion

5.1. Methode

Die hier verwendete Methode zur Bruchpunkt-Annäherung durch das Ausprobieren von Primer-Kombinationen ist einfach, schnell und liefert direkt ein Fusionsfragment zur Sequenzierung. Es wird dabei wenig DNA benötigt (Armour et al. 2002) und das Verfahren ist für solch kleine Introns mit maximal 23kb Länge, wie die des SPAST-Gens, gut geeignet und kosteneffektiv. Wenn man aber große Gene wie das DMD-Gen mit riesigen Introns (bis zu 250kb Größe) untersucht, werden Primeranzahlen in der Größenordnung von jeweils 200-300 forward- und reverse-Primer benötigt. Dies ist dann teuer und unhandlich, da nur wenig Automatisierung möglich ist. Teilweise gibt es Schwierigkeiten im Design, um als robustes Verfahren reproduzierbare Ergebnisse zu liefern (Armour et al. 2002). Außerdem dient die Methode nur der gezielten Suche in bestimmten Regionen, die schon eingegrenzt sein müssen. Da aber durch MLPA die betroffenen Introns schon bekannt waren, ist es hier als Methode gut geeignet. Ebenfalls wurden ähnliche Ansätze bereits unter anderem bei BRCA1-Deletionen sowie bei Deletionen und Duplikationen des LDL-Rezeptorgens erfolgreich eingesetzt (Payne et al. 2000, Kim et al. 1999).

Alternativ könnte man durch folgende Methoden bzw. deren Modifikation eine Bruchpunkt-Eingrenzung erreichen: CGH, SNP-Analyse, Southern Hybridization, MAPH und MLPA.

Das Prinzip bei CGH ist der quantitative Vergleich zweier unterschiedlich gefärbter Genome (zu untersuchendes Genom und Kontrolle), indem bestimmte Zielsequenzen kompetitiv co-hybridisiert werden (Bejjani und Shaffer 2006). Dabei werden zur Hybridisierung bei der klassischen CGH Metaphasen-Chromosome verwendet und bei der array-CGH Fragmente aus BAC-/PAC-Klonen (150kb lang), c-DNA oder synthetische Oligonukleotide (Shen et al. 2007). Bei der Verwendung von BAC-/PAC-Klonen müssen erst DNA-Fragmente aus diesen Klonen isoliert und PCR-amplifiziert werden (Carvalho et al. 2004). Die Hybridisierung wird quantifiziert durch das Überwiegen einer Farbe. Dies bedeutet, dass bestimmte Abschnitte bei der zu untersuchenden DNA über- bzw. unterrepräsentiert sind im Vergleich zur Kontrolle und somit Duplikationen und Deletionen entsprechen (Bejjani und Shaffer 2006). Damit kann man entweder das gesamte Genom auf CNAs (*copy number aberrations*) untersuchen oder man wählt eine kontinuierliche Abdeckung einer bestimmten Region. Die Auflösung ist dabei abhängig von der Größe der hybridisierenden Fragmente und deren Abstände. Somit limitiert die Nutzung von BAC-/PAC-Klonen die Auflösung auf maximal 80-200kb, was bei NF2-

Deletionen schon durchgeführt wurde, während c-DNA und Oligonukleotide noch nicht klinisch verwendet wurden (Armour et al. 2002). Dabei kann mit Oligonukleotiden theoretisch eine Auflösung von <100bp erreicht werden. Somit stellen oligo-array-CGH eine echte Weiterentwicklung dar. Sie sind kostengünstiger herzustellen als BAC-/PAC-Klone, leichter zu handhaben und schneller verfügbar durch das in silico-Design (Carvalho et al. 2004). So wurden auch oligo-array-CGH-Chips hergestellt (Shen et al. 2007). Diese dienen zwar der genomweiten Suche nach CNAs, abhängig von der Dichte der Oligonukleotide kann man aber in ganz bestimmten Regionen den Bruchpunkt schon sehr weit eingrenzen. Ein Vorteil eines Chips ist dabei die Reproduzierbarkeit und auch eine gewisse Flexibilität für Up-Dates, die mehr Informationen über die Zielregion enthalten (Shen et al. 2007). Dennoch ist der Einsatz von Chips sehr teuer und besonders bei Untersuchungen von kleinen Genen und bei einem geringen Untersuchungsumfang nicht effektiv.

Mittels SNP (*single nucleotide polymorphism*)-Analyse kann man auch auf CNAs hin untersuchen. Dabei erfolgt ein Restriktionsenzym-Verdau genomischer DNA, gefolgt von einer PCR-Amplifikation sowie der Auftrennung und Markierung (Zhao et al. 2004). Der nächste Schritt erfolgt durch die Hybridisierung mit einem SNP-array-Ansatz, welche durch Fluoreszenz visualisiert wird. Dabei liefert ein bestimmter Anteil an scheinbarer Homozygotie in einem bestimmten Bereich Hinweise für eine Deletion (Zhao et al. 2004). Es können mit einem 10K-SNP-Chip 10000 SNPs untersucht werden, was eine genomweite Auflösung von 150-300kb liefert (Zhao et al. 2004, Kloth et al. 2007). Es gibt aber auch schon 100K-SNP-Chips. Somit ist mit der Abdeckung einer hohen SNP-Dichte die Detektion von Mikrodeletions-Syndromen möglich (Zhao et al. 2004). Dieses Verfahren ist für eine genomweite Suche nach solchen Mutationen sinnvoll und effektiv. Durch eine Verfeinerung der Methode können auch einzelne Chromosomen und noch kleinere Regionen innerhalb eines Chromosoms untersucht werden (Zhao et al. 2004). Dafür wäre aber ein Extra-Chip notwendig und bei 2 Mio. SNPs/3Mrd. bp im Genom liegt die maximal zu erreichende Auflösung bei 1,5kb. Dies könnte zur Bruchpunkt-Eingrenzung schon ausreichen, wäre aber sehr teuer.

Southern Hybridization beruht auf einem Restriktionsenzym-Verdau einer genomischen Zielsequenz. Der Verdau wird dann aufgetrennt, geblottet und mit internen Sonden beprobt. Das Auftreten neuer bzw. der Verlust gewohnter Banden gibt dann Hinweise auf Aberrationen wie Deletionen oder Duplikationen. Dieses Verfahren ist auch semi-quantitativ möglich, indem die Intensität der Hybridisierungs-Banden mit einer Kontrolle verglichen werden, wobei eine stärkere Intensität einer Duplikation und eine geringere Intensität einer

Deletion entspricht (Sellner und Taylor 2004). Die Auflösung ist hierbei von der Dichte der Restriktionsstellen und der Auftrennung in der Elektrophorese abhängig. Somit ist es limitiert auf wenige gleichzeitig erfasste Loci, mit einer Dauer von mehreren Tagen sehr zeitaufwendig und es werden große DNA-Mengen benötigt (Sellner und Taylor 2004).

Bei MAPH wird genomische DNA auf einer Membran fixiert. Diese wird mit Proben aus Vektoren hybridisiert, welche gezielt unterschiedliche Größen besitzen (Sellner und Taylor 2004). Es folgt nach einem Waschschrift die PCR-Amplifikation der hybridisierten Fragmente, welche dann aufgetrennt werden. Die Intensität der Banden ergibt durch den Vergleich mit Kontrollen die Dosis bestimmter Regionen, also Deletionen und Duplikationen (Sellner und Taylor 2004). Dagegen erfolgt die Hybridisierung bei MLPA in Lösung. Es werden zwei Halbproben verwendet, die erst durch einen Ligations-Schritt eine kontinuierliche Probe, welche kleiner ist als bei MAPH, darstellen. Dabei besitzt eine der Halbproben eine variable Länge zur Auftrennung (Sellner und Taylor 2004).

Bei beiden ist das Klonieren von Fragmenten aus den Vektoren zeitaufwendig, wobei aber auch kommerzielle Sets für bestimmte Regionen/Gene existieren (Armour et al. 2002). Außerdem ist die Proben-Herstellung auch synthetisch möglich, wobei aber die Generierung variabler Längen schwierig ist (Sellner und Taylor 2004). So gibt es auch die Möglichkeit, Proben für ein „Fine-Mapping“ zu konstruieren, wobei die Proben intronische Sequenzen binden und Fragmente, statt über die Größe, über die Sequenz in einem array-Ansatz identifiziert werden. Deshalb wäre dafür auch MLPA hinsichtlich der Auflösung besser geeignet, da die Proben kleiner sind und man theoretisch bis in den Einzelbasen-Bereich detektieren könnte (Sellner und Taylor 2004, Armour et al. 2002). Allerdings wird noch kein Fusionsfragment mit solch einem Ansatz generiert und das Probendesign ist zeitaufwendig in dieser Größenordnung (Sellner und Taylor 2004).

Somit bestehen zwar alternative Methoden, es weisen aber alle Einschränkungen auf. Für den Probenumfang von zehn Templates, die eine ausgeprägte Heterogenität untereinander im Sinne von neun verschiedenen Exon-Kombinationen aufweisen und für einen forschungsbasierten Ansatz wie diese Arbeit, ist die verwendete Methode am besten geeignet.

5.2. PCR-Ergebnisse

5.2.1. Falsch positive PCR-Ergebnisse

Da insgesamt von allen zehn DNA-Templates Banden erhalten werden konnten, wurden von allen Templates PCR-Produkte weiterverarbeitet. Allerdings konnten erst nach der Sequenzierung und dem Vergleich mit der Referenz-Sequenz viele der Amplikons als unspezifisch identifiziert werden. Dies bedeutet, dass sie nicht den erwarteten Fusionssequenzen von SPAST-Deletionen entsprachen, sondern Sequenzen anderer genomischer Lokalisationen oder den Wildtyp darstellten.

Eine Ursache dafür könnte die zufällige Basengleichheit von intronischer SPAST-Sequenz mit einer Sequenz eines anderen Locus in der Größenordnung der Länge (um 24nt) der jeweiligen Primer sein. Bei einer Größe des menschlichen Genoms von 3×10^9 nt (Bentley 2000, Venter et al. 2001) ist dies zwar unwahrscheinlich, möglich wird diese potentielle Ursache aber durch die Tatsache, dass ein Primer zum Erkennen und Amplifizieren keine 100%ige Homologie zu seiner Zielsequenz benötigt. So konnten andere Bereiche im Genom durch die Primer erkannt und amplifiziert werden.

Eine weitere Möglichkeit zur Generierung unspezifischer PCR-Produkte ist das Binden der Primer an repetitive Sequenzelemente, die mehrfach im Genom vorkommen, obwohl die meisten dieser Strukturen beim Primerdesign ausgeschlossen wurden. Dabei machen insgesamt 33-42% des menschlichen Genoms eingestreute repetitive Sequenzen aus rudimentären, degenerierten, ehemals transposablen Elementen aus (Smit 1996, Smit 1999). Dies sind Strukturen, die sich über Zwischenverbindungen wie z.B. RNA-Transkripte aus ihrer ursprünglichen Lokalisation an anderen genomischen Stellen wieder einfügen können (Smit 1996). Die größte Gruppe dieser Elemente stellen Retrotransposons dar, wobei SINE ca. 12% und LINE 15% des Genoms ausmachen. Allerdings ist der SINE-/LINE-Gehalt vom GC-Gehalt abhängig. So kommen SINE eher in GC-reichen Regionen vor, während die LINE bevorzugt in GC-armen Regionen zu finden sind (Smit 1996, Henke et al. 1997, Smit 1999). Außer den Retrotransposons (SINE, LINE) und den DNA-Transposons (MER1, MER2) gibt es eine weitere, aber unklassifizierte Gruppe repetitiver transposabler Elemente, die zusammen 1% des menschlichen Genoms ausmachen (Smit 1996, Prak und Kazazian 2000). Weiterhin gibt es eine Unterform der Retrotransposons, die sich über RNA-Zwischenverbindungen von ihrer ursprünglichen Lokalisation an anderen Stellen einfügen, die als prozessierte Pseudogene Ähnlichkeit mit Alu-Elementen besitzen und 0,5% des

menschlichen Genoms darstellen (Prak und Kazazian 2000). Es können auch repetitive Elemente durch kleinere Deletionen, Duplikationen, Inversionen und weitere Insertionen stark degeneriert sein, so dass sie nicht mehr als solche erkannt werden können, aber gewisse Homologien zwischen den einzelnen Elementen bestehen (Smit 1996). So kommen DNA-Transposons zwar in einer geringen Anzahl an Kopien im menschlichen Genom vor und machen auch nur 1,5% davon aus, es existieren aber >150000 Sequenz-Homologe zu deletierten Mitgliedern dieser Gruppe (Smit 1996).

Eine weitere Quelle für eine nicht der gewünschten Fusion entsprechenden Bande stellte die Amplifikation des Wildtyps dar. So wurde bei P3 (siehe Abb. 4.8) sowohl der Wildtyp als auch die Fusionssequenz amplifiziert, während bei P9 außer dem Wildtyp keine Bande erhalten werden konnte. Eine mögliche Erklärung dafür ist die geringe Distanz zwischen den gewählten Primern auch schon beim Wildtyp.

5.2.2. Falsch negative PCR-Ergebnisse

Von zwei der zehn DNA-Templates konnte keine spezifische Bande erhalten werden, wobei bei P9 nur der Wildtyp amplifiziert werden konnte, während bei P4 nur unspezifische Sequenzen dargestellt werden konnten.

Möglicherweise wurden die Primer so gewählt, dass sie direkt auf dem Bruchpunkt gelegen hätten und die Distanz der dazu benachbarten Primer-Kombination für die PCR-Amplifikation zu lang gewesen wäre. Dafür wäre es nötig neue Primer zu bestellen, was aus finanziellen Gründen nicht realisiert wurde.

Ein anderer Grund dafür könnte sein, dass die Voraussetzungen für die PCR nicht optimal waren. So könnte sich ein lokal erhöhter GC-Gehalt der Zielsequenz und demzufolge auch der Primer erschwerend auf die Amplifikation auswirken (Henke et al. 1997). So spielt sowohl der GC-Gehalt des Templates als auch der der Primer-Kombination eine Rolle für den PCR-Erfolg, welcher mit einer Reduktion des GC-Gehalts steigt (Benita et al. 2003).

Dabei bewirkt ein lokal erhöhter GC-Gehalt eine Versteifung der Helix-Struktur und möglicherweise die Ausbildung von Haarnadelstrukturen, was die Denaturierung erschwert, lokal eine erhöhte Schmelztemperatur bewirkt und somit die Reaktion inhibiert (McDowell et al. 1998, Bachmann et al. 2003). Beispielhaft lässt sich der Locus des Fragilen-X-Syndroms durch eine Erhöhung von CGG-Repeats, was zu einem GC-Gehalt von >85% in diesem Bereich führt, nur schwer amplifizieren (Baskaran et al. 1996). Deshalb wäre eine

Optimierung hinsichtlich der Variationen in der Puffer-Konzentration, dem Puffer-pH, der Mg-Konzentration, der Denaturierungs- und Renaturierungszeit sowie der Anzahl der Zyklen notwendig (Sahdev et al. 2007). Außerdem existieren in der Literatur weitere Versuche und Strategien zur Optimierung der Reaktion und damit zur Verbesserung des PCR-Erfolgs. Diese beinhalten die Verwendung spezieller Taq-Polymerasen, Reagenzien-Zusatz bzw. Variation der Konzentration von z.B. Betaine, welches GC-Bindungen destabilisiert, eine Variierung der Temperatur und deren Dauer bei der Denaturierung und Renaturierung, eine längere Denaturierungsphase sowie das Einhalten eines bestimmten GC-Gehalts beim Primerdesign (Henke et al. 1997, McDowell et al. 1998, Baskaran et al. 1996, Sahdev et al. 2007, Bachmann et al. 2003). Insgesamt ist die Realisierung dieser Protokolle aber sehr arbeits- und zeitintensiv und war damit im Rahmen dieser Arbeit nicht möglich. Außerdem ist es auch nicht bekannt, ob ein lokal erhöhter GC-Gehalt verantwortlich ist für den Anteil falsch negativer PCR-Ergebnisse hier im Gegensatz zum Fragilen-X-Bereich (Baskaran et al. 1996). Dabei ist aufgrund der Bedeutung der Fragilen-X-Region eine Optimierung sinnvoller als für die Amplifikation einer einmaligen Fusionssequenz einer SPAST-Deletion.

Eine weitere Ursache, die das Template betrifft, besteht in der Möglichkeit, dass es sich bei P4, P9 oder bei beiden um komplexe Ereignisse handeln könnte, wobei zwar Exons fehlen, es sich aber nicht um einfache Deletionen handeln würde. So gibt es Publikationen von LIS1-Deletionen, bei denen ein Patient zusätzlich zu seiner partiellen LIS1-Deletion eine Insertion von 24bp zwischen angrenzendem 5'- und 3'-Bruchpunkt aufweist (Mei et al. 2008). Außerdem wurde bei einer DelEx8 im IDS-Gen (Hunter-Erkrankung) eine Insertion von 157bp im Bereich der Fusion gefunden, wobei ein großer Anteil der zusätzlichen Sequenz eine starke Homologie zu einem AluSg und einem L1-Element einer anderen Lokalisation aufweisen konnte. Der mögliche Mechanismus hier wurde als Alu-Retrotransposition mit nachfolgender Alu-vermittelter Deletion von 3kb erklärt (Ricci et al. 2003). Dabei führte die Retrotransposition eines AluSg zum Zusammenlagern zweier Alu-Elemente mit dem Resultat einer Deletion (Ricci et al. 2003). Außerdem wurde auch ein Fall einer 182bp-Insertion (zu Alu homologer Sequenz) im MSH2-Gen beschrieben, was aber keiner Retrotransposition entsprechen sollte, sondern eher eine Alu-vermittelte Rekombination darstellte, die begünstigt war durch einen hohen Alu-Gehalt in den angrenzenden Introns (Kloor et al. 2004). Allerdings ist die Größenordnung all dieser Insertionen zu klein, um das Misslingen der Amplifikationen zu erklären.

Dagegen beschrieben Lee et al. 2007 das Auftreten von 13/17 komplexen Duplikationen im PLP1-Gen bei nicht wiederkehrenden Ereignissen. Als komplex zu bezeichnen waren dabei

mehr als zwei Fusionen mit jeweiliger Mikrohomologie und zusätzlicher Sequenz im Bereich der Verbindungen. Es gab bei mehreren dieser Ereignisse einen Wechsel zwischen duplizierter DNA, triplizierten Anteilen, normaler DNA dazwischen und auch deletierte Bereiche in einer Größenordnung der jeweiligen Anteile mehrerer Kilobasen. Es wurde ein möglicher neuer Mechanismus dabei beschrieben, der nicht einer Rekombination während Mitose oder Meiose entspricht, sondern die Ereignisse durch fehlerhafte Replikationen erklärt. Dabei wird die normale Replikation verzögert oder angehalten und es kommt zum Wechsel der Replikationsgabel, die bis in Mb-Entfernung liegen kann. Der Wechsel wird durch das Vorhandensein von LCRs in der Nähe erleichtert und erfordert an den Fusionsstellen nur eine Mikrohomologie. Es kann auch nach einem erneuten Wechsel die ursprüngliche Sequenz weiter synthetisiert werden. So entstehen Deletionen, wenn Bereiche übersprungen werden, und Duplikationen, wenn die Replikation bereits synthetisierte Sequenz erfasst (Lee et al. 2007). Der Mechanismus kann damit komplexe Duplikationen und Deletionen erklären.

Komplexe Ereignisse mit einer partiellen Deletion und zusätzlichen Insertionen duplizierter Bereiche, z.B. von außerhalb des SPAST-Gens, welches eher klein ist, können bei P9 und P4 nicht ausgeschlossen werden. Eine eingefügte Duplikation von z.B. >10kb könnte eine Amplifikation verhindern. Um dies zu identifizieren wären Methoden wie Southern Hybridization oder CGH geeignet. Dies wäre aber zu zeitaufwendig gewesen.

5.3. Mechanismus und Struktur

Folgende Einteilung basiert auf der Homologie im Bereich der Fusionen der Bruchpunkte sowie in den 50bp- (in 5'- und 3'-Richtung) Bruchpunktumgebungen, da diese einen entscheidenden Hinweis auf den mutmaßlichen Mechanismus der Deletions-Generierung liefert.

5.3.1. Hinweise auf NHEJ

Sechs der acht hier charakterisierten SPAST-Deletionen sind gekennzeichnet durch eine Mikrohomologie am Fusionspunkt in einer Größenordnung von 3-7bp, was typisch für Deletionsereignisse ohne ausgeprägte Homologie ist (Roth und Wilson 1986). Eine ähnliche

Mikrohomologie wurde auch bei den meisten DMD-, IGHM-, TSC2-Deletionen, bei vielen PLP1-Duplikationen, einigen MSH2-Deletionen und einzelnen LIS1-, BTK- und NSD1-Deletionen gefunden (Toffolatti et al. 2002, Sironi et al. 2003, Kozlowski et al. 2007, Dabora et al. 2000, Woodward et al. 2005, Wagner et al. 2003, Mei et al. 2008, Douglas et al. 2005, van Zelm et al. 2008).

Zusätzlich wurden bei fast allen der erwähnten Gene auch mehrere Fälle mit Insertionen an der Deletionsverbindung von 2-15bp gefunden (Toffolatti et al. 2002, Sironi et al. 2003, Kozlowski et al. 2007, Dabora et al. 2000, Woodward et al. 2005, Wagner et al. 2003, van Zelm et al. 2008). Bei einer LIS1-Deletion wurde sogar eine 24bp lange Insertion gefunden (Mei et al. 2008). Wie die Insertionen im Bereich der Fusionen, sind auch kleine Duplikationen an dieser Lokalisation ein Merkmal für Ereignisse ohne ausgeprägte Homologie. Diese wurden ebenfalls schon bei mehreren DMD-Deletionen mit einer Größe von 9-24bp identifiziert (Toffolatti et al. 2002). Außerdem gibt es auch Deletionsereignisse ohne jegliche Übereinstimmung an der Fusion (Mei et al. 2008). Allerdings weisen die hier charakterisierten SPAST-Deletionen weder Insertionen, Duplikationen noch eine gänzlich fehlende Mikrohomologie auf. Eine mögliche Ursache dafür könnte die noch begrenzte Anzahl an SPAST-Deletionen mit Bruchpunkt-Analyse sein.

Alle Ereignisse mit einer Mikrohomologie weisen nur eine geringe Homologie in ihrer 100bp-Bruchpunktumgebung auf, die zwischen 21% und 34% liegt und damit um die zufällige Übereinstimmung schwankt.

Es sind davon auch drei Bruchpunkte in repetitiven Alu-Elementen lokalisiert, dabei sowohl der 5'- als auch der 3'-Bruchpunkt von P10. Allerdings kommen die beteiligten Elemente nicht in direkter Orientierung zueinander vor, sodass beim direkten Vergleich dieser Elemente keine ausgeprägte Homologie zwischen den Elementen festgestellt werden konnte. Ebenso reicht auch die Lokalisation einer der beiden Bruchpunkte in einem Alu-Element nicht aus für ein homologes Rekombinationsereignis (Sironi et al. 2003). Weiterhin weisen 33% der Bruchpunkte in nicht repetitiver Sequenz eine räumliche Nachbarschaft <50bp (siehe Tabelle 4.1) zu Alu-Elementen auf und weitere zwei Bruchpunkte (von derselben Deletion) befinden sich in der Nähe von LINE-Repeats. Allerdings liegt bei dem zuletzt genannten Ereignis keine ausgeprägte Homologie in der Bruchpunktumgebung und zwischen den beteiligten LINE-Repeats, die von verschiedenen Subfamilien abstammen, vor. So kann die Assoziation zu repetitiven Elementen bei fehlender Homologie insgesamt als Merkmal nicht homologer Rekombinationsereignisse aufgefasst werden. So fanden Sironi et al. 2003, Kozlowski et al. 2007, Ferec et al. 2006, Douglas et al. 2005, Wagner et al. 2003 und Mei et al. 2008 auch

Bruchpunkte in repetitiven Elementen oder in der Nähe ohne ausgeprägte Homologie. Das könnte mit den Instabilität verursachenden Effekten dieser Elemente erklärt werden (Shaw und Lupski 2004). Oder die bei SPAST-Deletionen gefundene Assoziation zu den repetitiven Elementen lässt sich durch deren hohen Gehalt erklären. So beträgt der Alu-Gehalt im SPAST-Gen ca. 40% (Beetz et al. 2006) und wenn die Doppelstrangbrüche unabhängig von Alu-vermittelten Ereignissen erfolgten, würden trotzdem statistisch 40% aller Bruchpunkte in Alu-Elementen lokalisiert sein. So liegen sechs der zwölf Bruchpunkte (50%) nicht homologer Deletionsereignisse sowohl innerhalb als auch in der Nähe eines Alu-Elements. Ähnliches kann über die Bruchpunkte von P8 mit der Nähe zu verschiedenen LINE-Sequenzen festgestellt werden. So ist der LINE-Gehalt in den angrenzenden Introns immens hoch mit 49% und 61% der Sequenz und damit die Wahrscheinlichkeit sehr groß, dass, wenn ein Bruchpunkt in diesen Introns gefunden wird, dieser auch innerhalb oder in der Nähe dieser Elemente lokalisiert ist. Es wurden aber auch homologe Deletionsereignisse zwischen LINE-Repeats beschrieben (Burwinkel und Kilimann 1998). Dafür ist allerdings eine ausgeprägte Homologie zwischen den Elementen die Voraussetzung. Auch sollten die Bruchpunkte direkt in den Elementen liegen. Insgesamt haben LINE eine geringere Bedeutung als Alu-Elemente für einen auf Homologie-basierten Mechanismus (Deininger und Batzer 1999).

Mit der Mikrohomologie in den Fusionsregionen, der fehlenden Homologie in den Bruchpunktumgebungen sowie der lediglich räumlichen Assoziation zu repetitiven Elementen können diese sechs Deletionsereignisse, also die Mehrheit, dem NHEJ als Mechanismus angerechnet werden.

5.3.2. Hinweise auf NAHR

Die anderen zwei hier analysierten SPAST-Deletionen weisen im Bereich der Deletionsverbindung eine höhere Übereinstimmung auf. Diese beträgt 14bp bei P2 und 6bp bei P5. Dabei ist die absolute Identität an der Fusion bei P5 zwar nur in der Größenordnung einer Mikrohomologie, bei Betrachtung der Homologie in der Bruchpunktumgebung weist P5 aber 63% Übereinstimmung auf, was wesentlich größer ist als die der anderen „Mikrohomologie-Ereignisse“. Diese Übereinstimmung ist sogar lokal noch erhöht auf 84% und entspricht damit der ebenfalls großen Homologie von P2 mit 83%. Beide Ereignisse weisen ihre Bruchpunkte an korrespondierender Stelle in direkt zueinander orientierten, 80%ig homologen Alu-Elementen auf, ohne dass es beim Vergleich der beiden eine

präferentielle Lokalisation im Alu-Element zu geben scheint (siehe auch Abb. 4.12). Solch eine bevorzugte Stelle innerhalb der Alu-Elemente wurde aber von Longa et al. 2001 bei drei durch NAHR verursachten TSC2-Deletionen beschrieben. Nur die lokale Erhöhung der Homologie bei P5 scheint eine Rekombination in diesem Bereich zu erleichtern, sodass sich der Bruch im Bereich höchster Übereinstimmung ereignet. Somit können diese beiden Ereignisse aufgrund der Homologie im Fusionsbereich, in der Bruchpunktumgebung und der Bruchpunkt-Lokalisation in homologen Alu-Elementen durch NAHR als Mechanismus erklärt werden.

Dabei ist sowohl die direkte Orientierung der beteiligten Elemente eine Voraussetzung, was auch bei homologen MSH2-, LIS1-, BRCA2- und NSD1-Deletionen so beschrieben wurde, als auch eine Homologie von über 80% (Kozak et al. 2006, Matejas et al. 2006, Wagner et al. 2003, Mei et al. 2008, Dabora et al. 2000, Agata et al. 2005, Douglas et al. 2005). Dies spiegelt sich auch in einer großen absoluten Homologie im Fusionsbereich wider. So liegt diese bei einer LIS1-Deletion bei 33bp, bei BRCA2 bis 34bp und bei SDHC zwischen 11 und 23bp (Mei et al. 2008, Agata et al. 2005, Baysal et al. 2004). Damit ist diese absolute Übereinstimmung sogar noch tendenziell größer als bei den hier charakterisierten homologen SPAST-Deletionen.

Eine noch größere Homologie als zwischen Alu-Elementen gibt es bei homologen Ereignissen, die sich zwischen LCRs abspielen. Die LCRs, die >100kb lang sind und Homologien >95% aufweisen, liefern im Fusionsbereich Übereinstimmungen bis 100bp und größer (Tatton-Brown et al. 2005, Reiter et al. 1996, Venturin et al. 2004, Shaw und Lupski 2005).

5.3.3. Dominanz eines Mechanismus

Es kommen bei großen (partiellen) SPAST-Deletionen beide Mechanismen (NHEJ und NAHR) vor, wobei die hier charakterisierten Bruchpunktumgebungen ein Überwiegen von nicht homologen Ereignissen vermuten lassen. Die einzige bisher veröffentlichte partielle SPG4-Duplikation (exon10_12dup), die in einer brasilianischen HSP-Familie identifiziert wurde und die durch eine Leserahmenverschiebung zu einem vorzeitigen Stop-Codon im zweiten Exon10 führt, weist in den Fusionsbereichen auch keine ausgeprägte Homologie auf, was für einen NHEJ-basierten Mechanismus in diesem Fall spricht (Mitne-Neto et al. 2007).

Von einer NHEJ-Dominanz kann aber hier noch nicht gesprochen werden, da dafür bisher zu wenige SPAST-Deletionsereignisse charakterisiert wurden.

Eine NHEJ-Dominanz bereits gut charakterisierter Gene mit großen Deletionen konnte bei DMD gezeigt werden, bei dessen Deletionen NAHR stark unterrepräsentiert ist (Toffolatti et al. 2002, McNaughton et al. 1998, Nobile et al. 2002, Sironi et al. 2003). Ebenfalls weisen schon relativ viele TSC2-Deletionen Merkmale auf, die auf NHEJ schließen lassen, während NAHR seltener vorkommt, genauso wie die Mehrheit (76%) der 33 charakterisierten MLH1-Deletionen auf NHEJ zurückzuführen ist (Kozlowski et al. 2007, Longa et al. 2001, Li et al. 2006). Außerdem sind alle 21 bereits analysierten CFTR-Deletionen auf NHEJ zurückzuführen, allerdings verfügt das CFTR-Gen auch nur über einen geringen Alu-Gehalt (Ferec et al. 2006). Weitere Gene mit bisherigem NHEJ-Überwiegen, aber geringerer Fallzahl sind IGHM und TSC1 (van Zelm et al. 2008, Longa et al. 2001).

Auch Gene mit Deletionsereignissen, die eine NAHR-Dominanz bei einer Fallzahl >20 aufweisen, sind bekannt: MSH2 und BRCA1. Dabei sind sowohl MSH2 (34,2%) als auch BRCA1 (42%) Gene mit einem hohen Alu-Gehalt (Li et al. 2006, Woodward et al. 2005a). Die meisten der Deletionen sind Alu-NAHR-vermittelt (Wagner et al. 2003, Charbonnier et al. 2005, Li et al. 2006, Nakagawa et al. 2003). Auch bei den bisher beschriebenen BTK- und Artemis-Gen-Deletionen scheint NAHR zu überwiegen (van Zelm et al. 2008).

Es wurden aber auch Gendeletionen, die durch LCR-NAHR vermittelt werden, beschrieben. So werden 84% der SMS-Deletionen durch SMS-LCRs vermittelt und 87% der NF1-Deletionen folgen diesem Mechanismus, wobei untypische NF1-Deletionen sowohl durch Alu-NAHR als auch durch NHEJ generiert werden sollten (Bi et al. 2003, Venturin et al. 2004). Auch bei CMT1A-Duplikationen und den rezibroken PMP22-Deletionen werden die meisten Rekombinationen durch NAHR zwischen den CMT1A-LCRs verursacht (Reiter et al. 1996).

Weiterhin ist dieser Mechanismus zwischen LCRs verantwortlich für viele Mikrodeletions-Syndrome wie Prader-Willi-, Angelmann- und das DiGeorge-Syndrom (Shaw und Lupski 2004).

Dieser Mechanismus ist bei SPAST nicht wahrscheinlich, da hier keine LCRs bekannt sind und die partiellen SPAST-Deletionen und ihre Bruchpunkte verschiedene Lokalisationen betreffen und weit verteilt sind (siehe 4.1.2).

Während die Rolle der Alus und LCRs bei der Beeinflussung der Chromatinstruktur und-ausrichtung sowie auf lokaler Ebene die Festlegung des Mechanismus im Sinne einer NAHR-Vermittlung unumstritten ist (Li et al. 2006), ist man auf der Suche nach

Rekombinationsmotiven, Strukturelementen und dem präferentiellen Auftreten bestimmter Nukleotide, die für NHEJ-Ereignisse prädisponieren. So wurden verschiedene Sequenzmotive in Assoziation mit großen Deletionen identifiziert. Toffolatti et al., McNaughton et al. und Sironi et al. fanden MAR-Motive in der Bruchpunktumgebung von DMD-Deletionen (Toffolatti et al. 2002, McNaughton et al. 1998, Sironi et al. 2003). Dieses Motiv soll eine Rolle spielen bei der DNA-Schleifenbildung. Auch im DMD-Intron7, welches viele Bruchpunkte enthält, wurden solche Motive gefunden, aber ohne Assoziation zu den Bruchpunkten (McNaughton et al. 1997). Es wurden weiterhin Topoisomerase-I-Schnittstellen sowohl bei partiellen PLP1-Duplikationen, welche aber nicht überrepräsentiert waren, als auch bei vielen DMD-Deletionen in der Bruchpunktumgebung identifiziert (Woodward et al. 2005b, Sironi et al. 2003). Topoisomerase-II-Schnittstellen wurden bei TSC2-Deletionen und bei nicht wiederkehrenden NF1-Deletionen gefunden (Dabora et al. 2000, Venturin et al. 2004). Weiterhin konnten *Immunoglobulin-Heavy-Chain-Class-Switch-Repeat*-Motive bei der Analyse von CFTR-Deletionen und bei untypischen NF1-Deletionen gefunden werden (Ferec et al. 2006, Venturin et al. 2004). Außerdem wurde in einem betroffenen DMD-Intron eine Translin-Bindestelle identifiziert, die aber nicht in Assoziation zu Bruchpunkten stand (Sironi et al. 2003). Daneben wurden auch bei homologen Ereignissen Motive gefunden, was gegen eine reine Prädisposition für NHEJ spricht. So wurde bei einer partiellen PMP22-Deletion ein *chi-like*-Sequenz-Motiv beschrieben (Matejas et al. 2006). Ebenso sind die Bruchpunkte von 14/15 MSH2-Deletionen in Alu-Kern-Sequenzen lokalisiert, welche teilweise aus chi-Sequenzen bestehen (Charbonnier et al. 2005). Allerdings konnten die Motive nicht zur Mechanismen-Aufklärung beitragen. Außerdem wurden bei der Analyse von >500 bekannten Rekombinationen von Genen bei Erbkrankheiten und Neoplasien nur die *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats* und die Translin-Bindestelle CTGCAT als überrepräsentierte Motive der Bruchpunktumgebungen bei Deletionsereignissen identifiziert (Abeyasinghe et al. 2003). So kommen, wie bereits erwähnt, auch z.B. Topoisomerase-Schnittstellen-Motive vor, sind aber nicht automatisch direkt involviert (Abeyasinghe et al. 2003). Somit wurde die Untersuchung auf Motive auf die *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats* und die Translin-Bindestellen (Translin spielt als Protein eine Rolle bei der Erkennung von Einzelstrang-DNA (Abeyasinghe et al. 2003)) in dieser Arbeit beschränkt.

Es wurden in den acht hier untersuchten Fusionssequenzen sechs *Immunoglobulin-Heavy-Chain-Class-Switch-Repeats* gefunden. Da allerdings in der Vergleichssequenz schon 13 Motive identifiziert werden konnten und im gesamten SPAST-Gen 545, müssten hier im

Mittel zehn gefunden werden und somit schwanken die Fusionssequenzen und die Vergleichssequenzen nur um diesen Wert. Von einer Anreicherung der Bruchpunktumgebungen mit diesen Motiven kann nicht ausgegangen werden. Zusätzlich wurde eine signifikante Anreicherung bei Abeyasinghe et al. auch nur in der 15bp-Umgebung gefunden, während die Motive hier bis 40bp Entfernung von den Bruchpunkten lokalisiert sind. Es wurde außerdem eine Translin-Bindestelle bei P2 gefunden. Es handelt sich bei dieser aber um GCCCAGCT, also nicht um die der überrepräsentierten von Abeyasinghe et al. Sie wurde nur einmal gefunden, also kann keine Entscheidung über eine mögliche Assoziation getroffen werden. Da es sich bei P2 um eine Alu-vermittelte Deletion handelt, trägt dieses gefundene Motiv auch nicht zur Aufklärung von NHEJ-Motiven bei.

Auch Strukturelemente wie Palindrome (invertierte Repeats), die fähig sein sollen Haarnadeln zu bilden und damit eine Sekundärstruktur-Bildung erleichtern, und alternierende Purin-, Pyrimidin-Strecken sollen laut Chuzhanova et al. mit Deletionsereignissen assoziiert sein (Chuzhanova et al. 2003). Allerdings entspricht die Anzahl und Länge der Palindrome in der Umgebung von SPAST-Deletions-Bruchpunkten der der zufällig ausgewählten Sequenzen.

Es konnten aber auch bei DMD-Deletionen einige Palindrome mit einer Größe von ≥ 6 bp eines Repeats und bei PLP1-Duplikationen ≥ 10 bp sowie die Sequenz AAATTT, die die DNA krümmen kann, bei PAH-Deletionen und DMD-Deletionen gefunden werden (McNaughton et al. 1998, Woodward et al. 2005b, Kozak et al. 2006, Toffolatti et al. 2002). Auch hierbei fehlt ein klarer Zusammenhang zum Mechanismus. Alternierende Purin-Pyrimidin-Strecken von bis zu 74bp (Chuzhanova et al. 2003) konnten in dieser Größenordnung in den SPAST-Deletions-Fusionen nicht gefunden werden. Die längste Strecke war nur 11bp lang. Solche Strecken wurden aber auch bei PLP1-Duplikationen gefunden (Woodward et al. 2005b). Weiterhin wurden Polypurin-Strecken bei CFTR-Deletionen und eine Region mit 2kb Polypurinen/-pyrimidinen bei PKD1-Deletionen identifiziert, die dort aber nicht assoziiert sind zu den Bruchpunkten, sondern eher eine global destabilisierende Rolle besitzen sollen (Ferec et al. 2006, Kozlowski et al. 2008).

Zusätzlich zu den Motiven und Strukturelementen soll ein AT-Reichtum im Bereich von Deletionsereignissen überrepräsentiert sein (Abeyasinghe et al. 2003). Der AT-Gehalt in den SPAST-Bruchpunktumgebungen scheint zwar mit 61% im Mittel erhöht, liegt aber im gesamten SPAST-Gen bei 60% und stellt damit auch keine Anreicherung in den Bruchpunktumgebungen dar. Es wurden zwar Poly-A/T-Strecken bei homologen TSC2-Deletionen berichtet sowie ein AT-Reichtum um den Hotspot in CMT1A-LCRs gefunden,

diese erklären jedoch keinen Zusammenhang zu NHEJ (Dabora et al. 2000, Reiter et al. 1996).

Somit scheinen Doppelstrangbrüche, die durch NHEJ repariert werden, sich im SPAST-Gen zufällig zu ereignen und nicht durch Motive und Strukturen vorhersagbar zu sein.

5.4. Bruchpunkt-Hotspot

Bei der theoretischen Analyse großer SPAST-Deletionen konnten keine Hinweise auf einen Hotspot gefunden werden. Es handelt sich um sehr heterogene Deletionsereignisse. Dabei besteht kein einheitliches Deletions-Ausmaß, wie es dagegen bei wiederkehrenden Deletionsereignissen der Fall ist. So macht die 1,5Mb große, wiederkehrende NF1-Deletion 80% aller NF1-Deletionen aus (Dorschner et al. 2000). 80% aller SMS-Patienten tragen die wiederkehrende 4Mb-Deletion (Chen et al. 1997). Auch die 2Mb-Deletion beim Sotos-Syndrom ist häufig (Rio et al. 2003). Dabei handelt es sich aber bei allen um Mikrodeletions-Syndrome zwischen LCRs, genauso wie das DiGeorge-Syndrom, Angelman-/Prader-Willi-Syndrom und das Williams-Beuren-Syndrom, welche jeweils gemeinsame Deletions-Größen zwischen LCRs besitzen (Shaikh et al. 2000, Amos-Landgraf et al. 1999, Valero et al. 2000). Dagegen variiert die Deletions-Größe bei den acht charakterisierten SPAST-Fusionen zwischen 700bp und 15000bp. Zusätzlich zeigte die Betrachtung aller großen SPAST-Deletionen, von denen die charakterisierten einen repräsentativen Teil darstellen, sowohl Einzel-Exon-Deletionen, Exon-Kombinationen als auch Deletionen des gesamten Gens, ohne dass eine Kombination dominiert (siehe 4.1.1.). Dagegen konnten bei sieben gleichen Exon-Kombinationen bei MSH2-Deletionen sogar die identischen Bruchpunkte identifiziert werden, was für die Möglichkeit von wiederkehrenden Alu-Rekombinationsereignissen spricht (Wagner et al. 2003). Bei Cystinose-Deletionen weisen sogar 1/3 aller Patienten eine homozygote, wiederkehrende 65kb-Deletion mit identischen Bruchpunkten auf (Forestier et al. 1999). Dagegen handelt es sich bei vier partiellen NSD1-Deletionen, die die gleiche Exon-Kombination aufweisen, um einmalige Deletionen mit unterschiedlichen Bruchpunkten (Douglas et al. 2005).

Des Weiteren sind die Bruchpunkte aller bekannten SPAST-Deletionen bezüglich der Lokalisation in verschiedenen Introns weit verteilt. Dabei kommen in fast allen Introns Bruchpunkte vor. Auch wenn Abb. 4.3 eine Bruchpunkt-Anreicherung im Bereich von Intron 1 und 16 sowie im extragenischen Bereich vermuten lässt, so berücksichtigt diese nicht die

unterschiedlichen Intron-Größen und somit nicht die Bruchpunktdichte. Diese ist aber ausschlaggebend für einen Hotspot, ohne ihn allgemeingültig definieren zu wollen, mit einer Größenordnung von ca. 80% der Bruchpunkte in ca. 10-20% der Sequenz (Myers et al. 2005). Das würde bei einer Größe des SPAST-Gens von ca. 90000bp eine Sequenz von z.B. 9000bp ausmachen. Dies entspricht den größeren SPAST-Introns. Wenn in einem dieser Introns 80% der Bruchpunkte lägen, würde dies bei 63 Ereignissen 50 Bruchpunkten entsprechen. Dagegen liegen aber z.B. im Intron7 (10735bp), welches zu den Bruchpunkt-reicheren Introns zählt und dieser Größe entspricht, nur zehn Bruchpunkte (siehe Abb. 4.3). Stattdessen gibt es bei den bekannten SPAST-Deletionen eine gute Korrelation zwischen Bruchpunkt-Anzahl und Intron-Größe (Abb. 4.4), aber keinen Zusammenhang zwischen Alu-Gehalt pro Intron und dem Auftreten von Bruchpunkten, was einen weiteren Beleg für eine weite und damit zufällige Bruchpunkt-Verteilung darstellt. Eine ähnliche Korrelation konnte im vermeintlichen Minor-Hotspot von DMD-Deletionen identifiziert werden. Dabei weist diese Region uniform verteilte Bruchpunkte auf und es soll ein Zusammenhang zwischen Bruchpunkt-Gehalt und Intron-Größen bestehen (Nobile et al. 1997).

Auch das Überwiegen von NHEJ als Mechanismus bei den acht charakterisierten SPAST-Deletionen spricht gegen einen Hotspot, wenn man davon ausgehen kann, dass die Doppelstrangbrüche, die durch NHEJ repariert werden, sich zufällig ereignen (Shaw und Lupski 2004). Allerdings wurden bei DMD-Deletionen sowohl ein Major-(60% aller Bruchpunkte) als auch ein Minor-Hotspot beschrieben, wobei zumindest der Major-Hotspot eine echte bruchanfällige Region mit weiteren Bruchpunkt-Clustern innerhalb dieses Hotspots zu sein scheint, obwohl die Mehrheit aller DMD-Deletionen durch NHEJ generiert wurde (Toffolatti et al. 2002, McNaughton et al. 1998). Auch bei den 21 beschriebenen CFTR-Deletionen, die fast ausschließlich auf NHEJ zurückzuführen sind, sollen zwei Bruchpunkt-Cluster mit drei bzw. 16 Bruchpunkten existieren (Ferec et al. 2006). Schneider et al. beschrieben einen Hotspot bei SHOX-Deletionen, welcher in einem 5kb-Bereich 73% aller 5'-Bruchpunkte enthält, während die 3'-Bruchpunkte variieren und es sich auch nicht um homologe Ereignisse handelt. Stattdessen verfügt die erwähnte Region über Tandem-Repeats, welche 85% der Sequenz ausmachen und die Ausbildung von Sekundärstrukturen und somit Instabilität verursachen sollen (Schneider et al. 2005). Ein weiterer, aber fraglicher Hotspot bei drei durch NHEJ verursachten TSC1-Deletionen wurde berichtet, bei dem die sechs Bruchpunkte in nur zwei Introns lokalisiert sind und eine Assoziation zu einem Pyrimidin-Reichtum in einem der betroffenen Introns gefunden wurde (Yip et al. 1999, Longa et al. 2001).

Somit scheinen Bruchpunkt-Hotspots nicht zwingend an homologe Rekombinationsereignisse gebunden zu sein. Dennoch ist ein Bruchpunkt-Hotspot bei SPAST-Deletionen unwahrscheinlich, kann aber nicht ausgeschlossen werden, da bisher nichts über die Verteilung der immerhin 31% extragenischen Bruchpunkte bekannt ist.

Dagegen sind bei wiederkehrenden Deletionsereignissen zwischen LCRs solche Hotspots typisch (Shaw und Lupski 2004). So gibt es bei den typischen NF1-Mikrodeletionen einen 2kb großen Hotspot innerhalb der NF1-LCRs, wobei außer der Homologie der LCRs auch *chi-like*-Elemente in beiden Repeats identifiziert wurden (Lopez-Correa et al. 2001). Auch die Mehrheit der HNPP-Deletions-Bruchpunkte liegt mit einer 53-fachen Häufung in einem 1,7kb großen Hotspot innerhalb von CMT1A-LCRs, wobei auch hier die Homologie allein nicht den Hotspot erklären soll. Dort wurde ein mariner-Transposon-like-Element zusätzlich in der Hotspot-Umgebung identifiziert (Reiter et al. 1996). Ähnlich verteilt liegen die Bruchpunkte auch bei SMS- und Sotos-Mikrodeletionen vor (Bi et al. 2003, Chen et al. 1997, Rio et al. 2003).

Aber auch bei nicht wiederkehrenden Deletionsereignissen sind Hotspots zu finden. Dabei wurden bei Alu-NAHR dominierten MSH2-Deletionen mehrere Cluster im Bereich von Intron1-7 beschrieben, ebenso zwei Cluster bei den NHEJ-dominierten CFTR-Deletionen (Wagner et al. 2003, Ferec et al. 2006). Dazu zählt weiterhin der DMD-Major-Hotspot und ein Cluster allerdings nur im distalen LCR von PMD-Duplikationen, wobei diese Ereignisse auch als nicht durch NAHR vermittelt beschrieben wurden (Toffolatti et al. 2002, Sironi et al. 2003, Nobile et al. 1997, Woodward et al. 2005b).

Außerdem spielen Rekombinations-Hotspots auch eine wesentliche Rolle bei allelischen Rekombinationsereignissen. So existiert ein 1,2kb langer Hotspot im TAP2-Gen (MHC II), in dem 81% der Bruchpunkte lokalisiert sind (Jeffreys et al. 2000). Ähnlich gibt es im Humanen- β -Globin-Gen einen 2kb großen Hotspot, der 75% der Bruchpunkte beherbergt. Dabei wurden in diesem Bereich ein Pur-Element sowie ein *chi-like*-Element gefunden (Smith et al. 1998). Ebenfalls gibt es einen Hotspot für Polymorphismus-Rekombinationen ohne Krankheitswert im PGM1-Gen, bei dem auch ein Pur-Element gefunden wurde, wobei die Rolle dieser Motive noch unklar ist (Yip et al. 1999).

Da bei den bekannten SPAST-Deletionen aber kein Hotspot gefunden wurde, kann kein allgemeiner PCR-Ansatz verwendet werden, um möglichst viele Deletionen aufdecken zu können. Stattdessen muss MLPA für das gesamte Gen durchgeführt werden.

6. Schlussfolgerung

Die hier vorgestellten Analysen bezüglich aller bekannten SPAST-Deletionen zeigen eine erstaunliche Vielfalt in den Kombinationen betroffener Exons. Außerdem konnte eine relativ homogene Verteilung der Deletions-Bruchpunkte im Gen gezeigt werden. Diese Befunde sprechen gegen das Vorliegen intragenischer Hotspot-Regionen. Eine entsprechende Aussage auch für extragenische Bruchpunkte erfordert weitere spezifische Untersuchungen.

Von acht der anfänglich ausgewählten zehn Proben konnte im Rahmen der Arbeit die Fusionssequenz ermittelt werden. Aus diesen Sequenzen konnte abgeleitet werden, dass NHEJ sechs Aberrationen zu Grunde liegt, während zwei Aberrationen auf NAHR zwischen homologen Alu-Sequenzbereichen zurückzuführen sind. Die Nicht-Amplifizierbarkeit der Fusion in zwei Proben mag als Hinweis auf die Relevanz weiterer Mutationsmechanismen, etwa komplexer Deletions-Insertions-Ereignisse, gelten.

Wie einige andere Gene, die häufig von pathogenen Deletionen betroffen sind, hat auch SPAST einen relativ hohen Gehalt an Alu-Sequenzen. Im Gegensatz zu vielen dieser Gene scheint das aber nicht mit einer Dominanz von NAHR verknüpft zu sein. Stattdessen legen die hier generierten Sequenzdaten nahe, dass NHEJ dominiert. Das wäre in Übereinstimmung zu DMD, dem wohl prominentesten Gen mit häufigen Deletionen.

Insgesamt bestätigt sich somit ein schon von „kleinen“ SPAST-Mutationen bekanntes Phänomen: auch SPAST-Deletionen können überall im Gen auftreten. Aus diesem Grund ist auch für den zugehörigen diagnostischen Schritt der Einschluss des gesamten Gens unerlässlich.

III. Literaturverzeichnis

- Abeyasinghe SS, Chuzhanova N, Krawczak M, Ball EV, Cooper DN. 2003. Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs. *Hum Mutat*, 22 (3):229-244.
- Agata S, Dalla Palma M, Callegaro M, Scaini MC, Menin C, Ghiotto C, Nicoletto O, Zavagno G, Chieco-Bianchi L, D'Andrea E, Montagna M. 2005. Large genomic deletions inactivate the BRCA2 gene in breast cancer families. *J Med Genet*, 42 (10):e64.
- Amos-Landgraf JM, Ji Y, Gottlieb W, Depinet T, Wandstrat AE, Cassidy SB, Driscoll DJ, Rogan PK, Schwartz S, Nicholls RD. 1999. Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am J Hum Genet*, 65 (2):370-386.
- Armour JA, Barton DE, Cockburn DJ, Taylor GR. 2002. The detection of large deletions or duplications in genomic DNA. *Hum Mutat*, 20 (5):325-337.
- Bachmann HS, Siffert W, Frey UH. 2003. Successful amplification of extremely GC-rich promoter regions using a novel 'slowdown PCR' technique. *Pharmacogenetics*, 13 (12):759-766.
- Barnes DE. 2001. Non-homologous end joining as a mechanism of DNA repair. *Curr Biol*, 11 (12):R455-457.
- Baskaran N, Kandpal RP, Bhargava AK, Glynn MW, Bale A, Weissman SM. 1996. Uniform amplification of a mixture of deoxyribonucleic acids with varying GC content. *Genome Res*, 6 (7):633-638.
- Baysal BE, Willett-Brozick JE, Filho PA, Lawrence EC, Myers EN, Ferrell RE. 2004. An Alu-mediated partial SDHC deletion causes familial and sporadic paraganglioma. *J Med Genet*, 41 (9):703-709.
- Beetz C, Nygren AO, Schickel J, Auer-Grumbach M, Burk K, Heide G, Kassubek J, Klimpe S, Klopstock T, Kreuz F, Otto S, Schule R, Schols L, Sperfeld AD, Witte OW, Deufel T. 2006. High frequency of partial SPAST deletions in autosomal dominant hereditary spastic paraplegia. *Neurology*, 67 (11):1926-1930.
- Bejjani BA, Shaffer LG. 2006. Application of array-based comparative genomic hybridization to clinical diagnostics. *J Mol Diagn*, 8 (5):528-533.

- Benita Y, Oosting RS, Lok MC, Wise MJ, Humphery-Smith I. 2003. Regionalized GC content of template DNA as a predictor of PCR success. *Nucleic Acids Res*, 31 (16):e99.
- Bentley DR. 2000. The Human Genome Project--an overview. *Med Res Rev*, 20 (3):189-196.
- Bi W, Park SS, Shaw CJ, Withers MA, Patel PI, Lupski JR. 2003. Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion or duplication of chromosome 17p11.2. *Am J Hum Genet*, 73 (6):1302-1315.
- Burwinkel B, Kilimann MW. 1998. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J Mol Biol*, 277 (3):513-517.
- Carvalho B, Ouwerkerk E, Meijer GA, Ylstra B. 2004. High resolution microarray comparative genomic hybridisation analysis using spotted oligonucleotides. *J Clin Pathol*, 57 (6):644-646.
- Charbonnier F, Baert-Desurmont S, Liang P, Di Fiore F, Martin C, Frerot S, Olschwang S, Wang Q, Buisine MP, Gilbert B, Nilbert M, Lindblom A, Frebourg T. 2005. The 5' region of the MSH2 gene involved in hereditary non-polyposis colorectal cancer contains a high density of recombinogenic sequences. *Hum Mutat*, 26 (3):255-261.
- Chen KS, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, Lupski JR. 1997. Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet*, 17 (2):154-163.
- Chuzhanova N, Abeysinghe SS, Krawczak M, Cooper DN. 2003. Translocation and gross deletion breakpoints in human inherited disease and cancer II: Potential involvement of repetitive sequence elements in secondary structure formation between DNA ends. *Hum Mutat*, 22 (3):245-251.
- Dabora SL, Nieto AA, Franz D, Jozwiak S, Van Den Ouweland A, Kwiatkowski DJ. 2000. Characterisation of six large deletions in TSC2 identified using long range PCR suggests diverse mechanisms including Alu mediated recombination. *J Med Genet*, 37 (11):877-883.
- Deininger PL, Batzer MA. 1999. Alu repeats and human disease. *Mol Genet Metab*, 67 (3):183-193.

- Depienne C, Tallaksen C, Lephay JY, Bricka B, Poëa-Guyon S, Fontaine B, Labauge P, Brice A, Durr A. 2006. Spastin mutations are frequent in sporadic spastic paraparesis and their spectrum is different from that observed in familial cases. *J Med Genet*, 43 (3):259-265.
- Depienne C, Fedirko E, Forlani S, Cazeneuve C, Ribai P, Feki I, Tallaksen C, Nguyen K, Stankoff B, Ruberg M, Stevanin G, Durr A, Brice A. 2007. Exon deletions of SPG4 are a frequent cause of hereditary spastic paraplegia. *J Med Genet*, 44 (4):281-284.
- Dorschner MO, Sybert VP, Weaver M, Pletcher BA, Stephens K. 2000. NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. *Hum Mol Genet*, 9 (1):35-46.
- Douglas J, Tatton-Brown K, Coleman K, Guerrero S, Berg J, Cole TR, Fitzpatrick D, Gillerot Y, Hughes HE, Pilz D, Raymond FL, Temple IK, Irrthum A, Schouten JP, Rahman N. 2005. Partial NSD1 deletions cause 5% of Sotos syndrome and are readily identifiable by multiplex ligation dependent probe amplification. *J Med Genet*, 42 (9):e56.
- Erichsen AK, Inderhaug E, Mattingsdal M, Eiklid K, Tallaksen CM. 2007. Seven novel mutations and four exon deletions in a collection of Norwegian patients with SPG4 hereditary spastic paraplegia. *Eur J Neurol*, 14 (7):809-814.
- Errico A, Ballabio A, Rugarli EI. 2002. Spastin, the protein mutated in autosomal dominant hereditary spastic paraplegia, is involved in microtubule dynamics. *Hum Mol Genet*, 11 (2):153-163.
- Ferec C, Casals T, Chuzhanova N, Macek M, Jr., Bienvenu T, Holubova A, King C, McDevitt T, Castellani C, Farrell PM, Sheridan M, Pantaleo SJ, Loumi O, Messaoud T, Cuppens H, Torricelli F, Cutting GR, Williamson R, Ramos MJ, Pignatti PF, Raguene O, Cooper DN, Audrezet MP, Chen JM. 2006. Gross genomic rearrangements involving deletions in the CFTR gene: characterization of six new events from a large cohort of hitherto unidentified cystic fibrosis chromosomes and meta-analysis of the underlying mechanisms. *Eur J Hum Genet*, 14 (5):567-576.
- Fink JK. 2003. The hereditary spastic paraplegias: nine genes and counting. *Arch Neurol*, 60 (8):1045-1049.
- Fonknechten N, Mavel D, Byrne P, Davoine CS, Cruaud C, Bonsch D, Samson D, Coutinho P, Hutchinson M, McMonagle P, Burgunder JM, Tartaglione A, Heinzlef O, Feki I, Deufel T, Parfrey N, Brice A, Fontaine B, Prud'homme JF, Weissenbach J, Durr A, Hazan J. 2000. Spectrum of SPG4 mutations in autosomal dominant spastic paraplegia. *Hum Mol Genet*, 9 (4):637-644.

- Forestier L, Jean G, Attard M, Cherqui S, Lewis C, van't Hoff W, Broyer M, Town M, Antignac C. 1999. Molecular characterization of CTNS deletions in nephropathic cystinosis: development of a PCR-based detection assay. *Am J Hum Genet*, 65 (2):353-359.
- Hanein S, Durr A, Ribai P, Forlani S, Leutenegger AL, Nelson I, Babron MC, Elleuch N, Depienne C, Charon C, Brice A, Stevanin G. 2007. A novel locus for autosomal dominant "uncomplicated" hereditary spastic paraplegia maps to chromosome 8p21.1-q13.3. *Hum Genet*, 122 (3-4):261-273.
- Hanein S, Martin E, Boukhris A, Byrne P, Goizet C, Hamri A, Benomar A, Lossos A, Denora P, Fernandez J, Elleuch N, Forlani S, Durr A, Feki I, Hutchinson M, Santorelli FM, Mhiri C, Brice A, Stevanin G. 2008. Identification of the SPG15 gene, encoding spastizin, as a frequent cause of complicated autosomal-recessive spastic paraplegia, including Kjellin syndrome. *Am J Hum Genet*, 82 (4):992-1002.
- Harding AE. 1981. Hereditary "pure" spastic paraplegia: a clinical and genetic study of 22 families. *J Neurol Neurosurg Psychiatry*, 44 (10):871-883.
- Harding AE. 1983. Classification of the hereditary ataxias and paraplegias. *Lancet*, 1 (8334):1151-1155.
- Henke W, Herdel K, Jung K, Schnorr D, Loening SA. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res*, 25 (19):3957-3958.
- Iwanaga H, Tsujino A, Shirabe S, Eguchi H, Fukushima N, Niikawa N, Yoshiura K, Eguchi K. 2005. Large deletion involving the 5'-UTR in the spastin gene caused mild phenotype of autosomal dominant hereditary spastic paraplegia. *Am J Med Genet A*, 133A (1):13-17.
- Jeffreys AJ, Ritchie A, Neumann R. 2000. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum Mol Genet*, 9 (5):725-733.
- Keppen LD, Leppert MF, O'Connell P, Nakamura Y, Stauffer D, Lathrop M, Lalouel JM, White R. 1987. Etiological heterogeneity in X-linked spastic paraplegia. *Am J Hum Genet*, 41 (5):933-943.

- Kim SH, Bae JH, Chae JJ, Kim UK, Choe SJ, Namkoong Y, Kim HS, Park YB, Lee CC. 1999. Long-distance PCR-based screening for large rearrangements of the LDL receptor gene in Korean patients with familial hypercholesterolemia. *Clin Chem*, 45 (9):1424-1430.
- Kozak L, Hrabincova E, Kintz J, Horky O, Zapletalova P, Blahakova I, Mejstrik P, Prochazkova D. 2006. Identification and characterization of large deletions in the phenylalanine hydroxylase (PAH) gene by MLPA: evidence for both homologous and non-homologous mechanisms of rearrangement. *Mol Genet Metab*, 89 (4):300-309.
- Kozlowski P, Bissler J, Pei Y, Kwiatkowski DJ. 2008. Analysis of PKD1 for genomic deletion by multiplex ligation-dependent probe assay: absence of hot spots. *Genomics*, 91 (2):203-208.
- Kozlowski P, Roberts P, Dabora S, Franz D, Bissler J, Northrup H, Au KS, Lazarus R, Domanska-Pakiela D, Kotulska K, Jozwiak S, Kwiatkowski DJ. 2007. Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. *Hum Genet*, 121 (3-4):389-400.
- Krawczak M, Cooper DN. 1991. Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. *Hum Genet*, 86 (5):425-441.
- Li L, McVety S, Younan R, Liang P, Du Sart D, Gordon PH, Hutter P, Hogervorst FB, Chong G, Foulkes WD. 2006. Distinct patterns of germ-line deletions in MLH1 and MSH2: the implication of Alu repetitive element in the genetic etiology of Lynch syndrome (HNPCC). *Hum Mutat*, 27 (4):388.
- Longa L, Saluto A, Brusco A, Polidoro S, Padovan S, Allavena A, Carbonara C, Grosso E, Migone N. 2001. TSC1 and TSC2 deletions differ in size, preference for recombinatorial sequences, and location within the gene. *Hum Genet*, 108 (2):156-166.
- Lopez-Correa C, Dorschner M, Brems H, Lazaro C, Clementi M, Upadhyaya M, Dooijes D, Moog U, Kehrer-Sawatzki H, Rutkowski JL, Fryns JP, Marynen P, Stephens K, Legius E. 2001. Recombination hotspot in NF1 microdeletion patients. *Hum Mol Genet*, 10 (13):1387-1392.
- Lupski JR. 2004. Hotspots of homologous recombination in the human genome: not all homologous sequences are equal. *Genome Biol*, 5 (10):242.

- Mannan AU, Krawen P, Sauter SM, Boehm J, Chronowska A, Paulus W, Neesen J, Engel W. 2006. ZFYVE27 (SPG33), a novel spastin-binding protein, is mutated in hereditary spastic paraplegia. *Am J Hum Genet*, 79 (2):351-357.
- Matejas V, Huehne K, Thiel C, Sommer C, Jakubiczka S, Rautenstrauss B. 2006. Identification of Alu elements mediating a partial PMP22 deletion. *Neurogenetics*, 7 (2):119-126.
- McDermott C, White K, Bushby K, Shaw P. 2000. Hereditary spastic paraparesis: a review of new developments. *J Neurol Neurosurg Psychiatry*, 69 (2):150-160.
- McDermott CJ, Burness CE, Kirby J, Cox LE, Rao DG, Hewamadduma C, Sharrack B, Hadjivassiliou M, Chinnery PF, Dalton A, Shaw PJ. 2006. Clinical features of hereditary spastic paraplegia due to spastin mutation. *Neurology*, 67 (1):45-51.
- McDowell DG, Burns NA, Parkes HC. 1998. Localised sequence regions possessing high melting temperatures prevent the amplification of a DNA mimic in competitive PCR. *Nucleic Acids Res*, 26 (14):3340-3347.
- McMonagle P, Byrne PC, Fitzgerald B, Webb S, Parfrey NA, Hutchinson M. 2000. Phenotype of AD-HSP due to mutations in the SPAST gene: comparison with AD-HSP without mutations. *Neurology*, 55 (12):1794-1800.
- McNaughton JC, Hughes G, Jones WA, Stockwell PA, Klamut HJ, Petersen GB. 1997. The evolution of an intron: analysis of a long, deletion-prone intron in the human dystrophin gene. *Genomics*, 40 (2):294-304.
- McNaughton JC, Cockburn DJ, Hughes G, Jones WA, Laing NG, Ray PN, Stockwell PA, Petersen GB. 1998. Is gene deletion in eukaryotes sequence-dependent? A study of nine deletion junctions and nineteen other deletion breakpoints in intron 7 of the human dystrophin gene. *Gene*, 222 (1):41-51.
- Mei D, Lewis R, Parrini E, Lazarou LP, Marini C, Pilz DT, Guerrini R. 2008. High frequency of genomic deletions--and a duplication--in the LIS1 gene in lissencephaly: implications for molecular diagnosis. *J Med Genet*, 45 (6):355-361.
- Meijer IA, Hand CK, Cossette P, Figlewicz DA, Rouleau GA. 2002. Spectrum of SPG4 mutations in a large collection of North American families with hereditary spastic paraplegia. *Arch Neurol*, 59 (2):281-286.

- Mitne-Neto M, Kok F, Beetz C, Pessoa A, Bueno C, Graciani Z, Martyn M, Monteiro CB, Mitne G, Hubert P, Nygren AO, Valadares M, Cerqueira AM, Starling A, Deufel T, Zatz M. 2007. A multi-exonic SPG4 duplication underlies sex-dependent penetrance of hereditary spastic paraplegia in a large Brazilian pedigree. *Eur J Hum Genet*, 15 (12):1276-1279.
- Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of recombination rates and hotspots across the human genome. *Science*, 310 (5746):321-324.
- Nakagawa H, Hampel H, de la Chapelle A. 2003. Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. *Hum Mutat*, 22 (3):258.
- Nielsen JE, Koefoed P, Abell K, Hasholt L, Eiberg H, Fenger K, Niebuhr E, Sorensen SA. 1997. CAG repeat expansion in autosomal dominant pure spastic paraplegia linked to chromosome 2p21-p24. *Hum Mol Genet*, 6 (11):1811-1816.
- Nobile C, Marchi J, Nigro V, Roberts RG, Danieli GA. 1997. Exon-intron organization of the human dystrophin gene. *Genomics*, 45 (2):421-424.
- Nobile C, Toffolatti L, Rizzi F, Simionati B, Nigro V, Cardazzo B, Patarnello T, Valle G, Danieli GA. 2002. Analysis of 22 deletion breakpoints in dystrophin intron 49. *Hum Genet*, 110 (5):418-421.
- Pastink A, Eeken JC, Lohman PH. 2001. Genomic integrity and the repair of double-strand DNA breaks. *Mutat Res*, 480-481:37-50.
- Patrono C, Scarano V, Cricchi F, Melone MA, Chiriaco M, Napolitano A, Malandrini A, De Michele G, Petrozzi L, Giraldi C, Santoro L, Servidei S, Casali C, Filla A, Santorelli FM. 2005. Autosomal dominant hereditary spastic paraplegia: DHPLC-based mutation analysis of SPG4 reveals eleven novel mutations. *Hum Mutat*, 25 (5):506.
- Payne SR, Newman B, King MC. 2000. Complex germline rearrangement of BRCA1 associated with breast and ovarian cancer. *Genes Chromosomes Cancer*, 29 (1):58-62.
- Polo JM, Calleja J, Combarros O, Berciano J. 1991. Hereditary ataxias and paraplegias in Cantabria, Spain. An epidemiological and clinical study. *Brain*, 114 (Pt 2):855-866.

- Prak ET, Kazazian HH, Jr. 2000. Mobile elements and the human genome. *Nat Rev Genet*, 1 (2):134-144.
- Reid E. 1999. The hereditary spastic paraplegias. *J Neurol*, 246 (11):995-1003.
- Reid E. 2003. Many pathways lead to hereditary spastic paraplegia. *Lancet Neurol*, 2 (4):210.
- Reiter LT, Murakami T, Koeuth T, Pentao L, Muzny DM, Gibbs RA, Lupski JR. 1996. A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. *Nat Genet*, 12 (3):288-297.
- Ricci V, Regis S, Di Duca M, Filocamo M. 2003. An Alu-mediated rearrangement as cause of exon skipping in Hunter disease. *Hum Genet*, 112 (4):419-425.
- Rio M, Clech L, Amiel J, Faivre L, Lyonnet S, Le Merrer M, Odent S, Lacombe D, Edery P, Brauner R, Raoul O, Gosset P, Prieur M, Vekemans M, Munnich A, Colleaux L, Cormier-Daire V. 2003. Spectrum of NSD1 mutations in Sotos and Weaver syndromes. *J Med Genet*, 40 (6):436-440.
- Roth DB, Wilson JH. 1986. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol Cell Biol*, 6 (12):4295-4304.
- Roth DB, Porter TN, Wilson JH. 1985. Mechanisms of nonhomologous recombination in mammalian cells. *Mol Cell Biol*, 5 (10):2599-2607.
- Sahdev S, Saini S, Tiwari P, Saxena S, Singh Saini K. 2007. Amplification of GC-rich genes by following a combination strategy of primer design, enhancers and modified PCR cycle conditions. *Mol Cell Probes*, 21 (4):303-307.
- Sauter S, Mitterski B, Klimpe S, Bonsch D, Schols L, Visbeck A, Papke T, Hopf HC, Engel W, Deufel T, Epplen JT, Neesen J. 2002. Mutation analysis of the spastin gene (SPG4) in patients in Germany with autosomal dominant hereditary spastic paraplegia. *Hum Mutat*, 20 (2):127-132.
- Schneider KU, Sabherwal N, Jantz K, Roth R, Muncke N, Blum WF, Cutler GB, Jr., Rappold G. 2005. Identification of a major recombination hotspot in patients with short stature and SHOX deficiency. *Am J Hum Genet*, 77 (1):89-96.
- Sellner LN, Taylor GR. 2004. MLPA and MAPH: new techniques for detection of gene deletions. *Hum Mutat*, 23 (5):413-419.

- Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS. 2000. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet*, 9 (4):489-501.
- Shaw CJ, Lupski JR. 2004. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet*, 13 Spec No 1:R57-64.
- Shaw CJ, Lupski JR. 2005. Non-recurrent 17p11.2 deletions are generated by homologous and non-homologous mechanisms. *Hum Genet*, 116 (1-2):1-7.
- Shen Y, Irons M, Miller DT, Cheung SW, Lip V, Sheng X, Tomaszewicz K, Shao H, Fang H, Tang HS, Walsh CA, Platt O, Gusella JF, Wu BL. 2007. Development of a focused oligonucleotide-array comparative genomic hybridization chip for clinical diagnosis of genomic imbalance. *Clin Chem*, 53 (12):2051-2059.
- Sironi M, Pozzoli U, Cagliani R, Giorda R, Comi GP, Bardoni A, Menozzi G, Bresolin N. 2003. Relevance of sequence and structure elements for deletion events in the dystrophin gene major hot-spot. *Hum Genet*, 112 (3):272-288.
- Smit AF. 1996. The origin of interspersed repeats in the human genome. *Curr Opin Genet Dev*, 6 (6):743-748.
- Smit AF. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev*, 9 (6):657-663.
- Smith RA, Ho PJ, Clegg JB, Kidd JR, Thein SL. 1998. Recombination breakpoints in the human beta-globin gene cluster. *Blood*, 92 (11):4415-4421.
- Stankiewicz P, Lupski JR. 2002a. Genome architecture, rearrangements and genomic disorders. *Trends Genet*, 18 (2):74-82.
- Stankiewicz P, Lupski JR. 2002b. Molecular-evolutionary mechanisms for genomic disorders. *Curr Opin Genet Dev*, 12 (3):312-319.
- Starling A, Rocco P, Passos-Bueno MR, Hazan J, Marie SK, Zatz M. 2002. Autosomal dominant (AD) pure spastic paraplegia (HSP) linked to locus SPG4 affects almost exclusively males in a large pedigree. *J Med Genet*, 39 (12):e77.

- Stevanin G, Santorelli FM, Azzedine H, Coutinho P, Chomilier J, Denora PS, Martin E, Ouvrard-Hernandez AM, Tessa A, Bouslam N, Lossos A, Charles P, Loureiro JL, Elleuch N, Confavreux C, Cruz VT, Ruberg M, Leguern E, Grid D, Tazir M, Fontaine B, Filla A, Bertini E, Durr A, Brice A. 2007. Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. *Nat Genet*, 39 (3):366-372.
- Tatton-Brown K, Douglas J, Coleman K, Baujat G, Chandler K, Clarke A, Collins A, Davies S, Faravelli F, Firth H, Garrett C, Hughes H, Kerr B, Liebelt J, Reardon W, Schaefer GB, Splitt M, Temple IK, Waggoner D, Weaver DD, Wilson L, Cole T, Cormier-Daire V, Irrthum A, Rahman N. 2005. Multiple mechanisms are implicated in the generation of 5q35 microdeletions in Sotos syndrome. *J Med Genet*, 42 (4):307-313.
- Toffolatti L, Cardazzo B, Nobile C, Danieli GA, Gualandi F, Muntoni F, Abbs S, Zanetti P, Angelini C, Ferlini A, Fanin M, Patarnello T. 2002. Investigating the mechanism of chromosomal deletion: characterization of 39 deletion breakpoints in introns 47 and 48 of the human dystrophin gene. *Genomics*, 80 (5):523-530.
- Tsaousidou MK, Ouahchi K, Warner TT, Yang Y, Simpson MA, Laing NG, Wilkinson PA, Madrid RE, Patel H, Hentati F, Patton MA, Hentati A, Lamont PJ, Siddique T, Crosby AH. 2008. Sequence alterations within CYP7B1 implicate defective cholesterol homeostasis in motor-neuron degeneration. *Am J Hum Genet*, 82 (2):510-515.
- Tsukamoto Y, Ikeda H. 1998. Double-strand break repair mediated by DNA end-joining. *Genes Cells*, 3 (3):135-144.
- Valdmanis PN, Meijer IA, Reynolds A, Lei A, MacLeod P, Schlesinger D, Zatz M, Reid E, Dion PA, Drapeau P, Rouleau GA. 2007. Mutations in the KIAA0196 gene at the SPG8 locus cause hereditary spastic paraplegia. *Am J Hum Genet*, 80 (1):152-161.
- Valero MC, de Luis O, Cruces J, Perez Jurado LA. 2000. Fine-scale comparative mapping of the human 7q11.23 region and the orthologous region on mouse chromosome 5G: the low-copy repeats that flank the Williams-Beuren syndrome deletion arose at breakpoint sites of an evolutionary inversion(s). *Genomics*, 69 (1):1-13.
- van Zelm MC, Geertsema C, Nieuwenhuis N, de Ridder D, Conley ME, Schiff C, Tezcan I, Bernatowska E, Hartwig NG, Sanders EA, Litzman J, Kondratenko I, van Dongen JJ, van der Burg M. 2008. Gross deletions involving IGHM, BTK, or Artemis: a model for genomic lesions mediated by transposable elements. *Am J Hum Genet*, 82 (2):320-332.

- Venturin M, Gervasini C, Orzan F, Bentivegna A, Corrado L, Colapietro P, Friso A, Tenconi R, Upadhyaya M, Larizza L, Riva P. 2004. Evidence for non-homologous end joining and non-allelic homologous recombination in atypical NF1 microdeletions. *Hum Genet*, 115 (1):69-80.
- Wagner A, Barrows A, Wijnen JT, van der Klift H, Franken PF, Verkuijlen P, Nakagawa H, Geugien M, Jaghmohan-Changur S, Breukel C, Meijers-Heijboer H, Morreau H, van Puijenbroek M, Burn J, Coronel S, Kinarski Y, Okimoto R, Watson P, Lynch JF, de la Chapelle A, Lynch HT, Fodde R. 2003. Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the MSH2 gene. *Am J Hum Genet*, 72 (5):1088-1100.
- Woodward AM, Davis TA, Silva AG, Kirk JA, Leary JA. 2005a. Large genomic rearrangements of both BRCA2 and BRCA1 are a feature of the inherited breast/ovarian cancer phenotype in selected families. *J Med Genet*, 42 (5):e31.
- Woodward KJ, Cundall M, Sperle K, Sistermans EA, Ross M, Howell G, Gribble SM, Burford DC, Carter NP, Hobson DL, Garbern JY, Kamholz J, Heng H, Hodes ME, Malcolm S, Hobson GM. 2005b. Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. *Am J Hum Genet*, 77 (6):966-987.
- Yabe I, Sasaki H, Tashiro K, Matsuura T, Takegami T, Satoh T. 2002. Spastin gene mutation in Japanese with hereditary spastic paraplegia. *J Med Genet*, 39 (8):e46.
- Yip AG, Durr A, Marchuk DA, Ashley-Koch A, Hentati A, Rubinsztein DC, Reid E. 2003. Meta-analysis of age at onset in spastin-associated hereditary spastic paraplegia provides no evidence for a correlation with mutational class. *J Med Genet*, 40 (9):e106.
- Yip SP, Lovegrove JU, Rana NA, Hopkinson DA, Whitehouse DB. 1999. Mapping recombination hotspots in human phosphoglucomutase (PGM1). *Hum Mol Genet*, 8 (9):1699-1706.
- Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C, Gray JW, Sellers WR, Meyerson M. 2004. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res*, 64 (9):3060-3071.

IV. Danksagung

Ich möchte mich ganz herzlich bei der gesamten Arbeitsgruppe „Klinische Chemie“ des Instituts für Klinische Chemie und Laboratoriumsdiagnostik der FSU Jena für die Unterstützung bei der Anfertigung meiner Doktorarbeit bedanken.

Prof. Dr. Thomas Deufel gilt besonderer Dank für die Bereitstellung des Themas sowie für die Möglichkeit, den experimentellen Teil in seinem Labor unter professionellen Bedingungen durchführen zu können.

Bei Dr. Christian Beetz möchte ich mich bedanken für die Ratschläge und Hilfestellungen für das Gelingen der Versuche sowie für konstruktive Hinweise und Diskussion beim schriftlichen Teil der Arbeit.

Besonderer Dank gilt Kerstin Stein für die geduldige und freundliche Einarbeitung, Unterstützung und Zuarbeit bei den Experimenten.

Außerdem möchte ich meinem Mann Marouane, meinem Vater und meinem Bruder Stefan für die Hilfestellung und Überwindung einiger technischer Hindernisse beim schriftlichen Teil der Arbeit danken.

V. Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben:

Prof. Dr. Thomas Deufel

Dr. Christian Beetz,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Ort, Datum

Unterschrift des Verfassers

VI. Wissenschaftlicher Beitrag

Poster:

C. Beetz, C. Oubrayme, C. Depienne, S. Züchner, E. Reid, J. Schickel, L. Schöls, R. Schüle, M. Auer-Grumbach, S. Klebe, M. Pericak-Vance, A. Brice, T. Deufel. Comprehensive Characterization of Deletions in the SPAST Gene. *57th Annual Meeting of the American Society of Human Genetics*. 23.-27. Oktober 2007; San Diego, USA.